

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 17 Seconds
(without alignments)
2468.738 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEEPGATPQVILGLVLELR.....GLPPSPNEPATEHPPEQDET 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3923	92.7	804	1	MEAG6 HUMAN
2	276.5	6.5	1959	1	MYH9 CHICK
3	274.5	6.5	1130	1	YL17 CAEL
4	272	6.4	3210	1	CENF HUMAN
5	271.5	6.4	1875	1	MLP1 YEAST
6	270.5	6.4	1976	1	MYHA BOVIN
7	270	6.4	1357	1	KYNI HUMAN
8	268	6.3	1957	1	SPOF SCHPO
9	266.5	6.3	1790	1	MYHA HUMAN
10	266	6.3	1790	1	US01 YEAST
11	265	6.3	1411	1	EEA1 HUMAN
12	258.5	6.1	1976	1	MYHA RAT
13	258	6.1	2017	1	MYSN DROME
14	257	6.1	1583	1	GCC2 HUMAN
15	254	6.0	1679	1	CCC2 MOUSE
16	253	6.0	802	1	ENAH MOUSE
17	253	6.0	1938	1	BNI1 YEAST
18	253	6.0	1953	1	BN1 YEAST
19	252.5	6.0	944	1	NUPI YEAST
20	252	6.0	1935	1	MYAS CYPCA
21	251	5.9	997	1	SCPI RAT
22	251	5.9	1978	1	MYHE CHICK
23	251	5.9	3259	1	GOBI HUMAN
24	249.5	5.9	1020	1	CF60 HUMAN
25	249.5	5.9	1960	1	MYH9 HUMAN
26	249	5.9	1961	1	MYH9 RAT
27	247.5	5.8	1938	1	MYHD HUMAN
28	246.5	5.8	1979	1	TRIA HUMAN
29	245	5.8	1972	1	MYHB HUMAN
30	245	5.8	2871	1	DESP HUMAN
31	244.5	5.8	1935	1	MYH7 PIG
32	244.5	5.8	2116	1	MYSD DICDI
33	243.5	5.8	1935	1	MYH7 HUMAN

34 243 5.7 2611 1 BP1E MOUSE
35 242.5 5.7 1102 1 MYSC CHICK
36 242.5 5.7 1597 1 CTRO_MOUSE
37 241 5.7 1330 1 KINI_VULVU
38 241 5.7 1940 1 MYH3 RAT
39 240.5 5.7 697 1 MFPI LYCES
40 240.5 5.7 1690 1 C190 DROME
41 240 5.7 1972 1 MYHB MOUSE
42 240 5.7 2054 1 M18A HUMAN
43 239 5.6 1248 1 DIA1 HUMAN
44 239 5.6 1938 1 MYH6 MOUSE
45 238.5 5.6 1727 1 ALM1 SCHPO

ALIGNMENTS

RESULT 1
ID MEAG6 HUMAN STANDARD; PRT; 804 AA.
AC O15320: O00169;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Meningioma-expressed antigen 6/11 (MEAG6) (MEAL1).
GN MGEA6 OR MGEA11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS MEAG6 AND MEAL1).
RC TISSUE=Meningioma;
RX MEDLINE=97472454; PubMed=9356211;
RA Heckel D., Brass N., Fischer U., Blin N., Steudel I., Tuereci O.,
RA Packler O., Zang K.D., Meese E.;
RT "cDNA cloning and chromosomal mapping of a predicted coiled-coil
RT proline-rich protein immunogenic in meningioma patients.";
RL Hum. Mol. Genet. 6:2031-2041(1997).
CC -!- ALTERNATIVE PRODUCTS:
CC -!- Event=Alternative splicing; Named isoforms=2;
CC Name=MEAG6;
CC IsoId=O15320-1; Sequences=Displayed;
CC Name=MEAL1;
CC IsoId=O15320-2; Sequences=VSP 004322;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in brain, muscle
CC and cranial skin.
CC -!- DISEASE: AUTOCANTIGEN FOUND IN BENIGN MENINGIOMA.
CC -!- CAUTION: At least 9 pseudogenes have been found on different
CC chromosomes, including chromosomes 2, 3, 6, 7, 9, 10, 12, 13 and
CC 18.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U94780; AAB86593.1;
EMBL; U73682; AAB86589.1; ALT_INIT.
Genew; HGNC:7057; MGEA6.
MIM; 602132;
DR GO; GO:0008047; F:enzyme activator activity; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
KW Antigen; Coiled coil; Alternative splicing
FT DOMAIN 87 273 COILED COIL (POTENTIAL).
FT DOMAIN 321 501 COILED COIL (POTENTIAL).
FT DOMAIN 553 558 POLY-GLY.
FT DOMAIN 502 804 PRO-RICH.
FT VARSPLIC 514 556 Missing (in isoform MEAL1).

```

FT CONFLICT 198 198 /FTID=VSP 004322.
FT CONFLICT 360 360 P -> Q (IN REF. 1; AAB85889).
FT CONFLICT 589 589 L -> P (IN REF. 1; AAB85889).
FT CONFLICT 594 594 L -> S (IN REF. 1; AAB85889).
FT CONFLICT 699 699 G -> I (IN REF. 1; AAB85889).
FT CONFLICT 738 738 V -> R (IN REF. 1; AAB85889).
FT CONFLICT 804 AA; 30943 MW; 3F088799E0883EDE CRC64;
SQ SEQUENCE 804 AA; 30943 MW; 3F088799E0883EDE CRC64;

Query Match 92.7%; Score 3923; DB 1; Length 804;
Best Local Similarity 93.1%; Pred. No. 3e-136;
Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;

QY 1 MEEPGATPOYGLVLEELRRVVAALPEMGRDENPYGPPSELVCAAVGVFFVLLFLW 60
DB 1 MEEPGATPOYGLVLEELRRVVAALPEMGRDPSNLYGPPSELVCAAVGVFFVLLFLW 60

QY 61 RSFRSRSRLVYVREKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELDASFE 120
DB 61 RSFRSRSRLVYVREKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELDASFE 120

QY 121 AAEARSLEATCEKLNRSSELEDETLCKELKOKSKHQSQDELMADISKISQLEDE 180
DB 121 AAEARSLEATCEKLNRSSELEDETLCKELKOKSKHQSQDELMADISKISQLEDE 180

QY 121 AT-EAQSLEATCEKLNRSSELEDETLCKELKOKSKHQSQDELMADISKISQLEDE 179
DB 121 AT-EAQSLEATCEKLNRSSELEDETLCKELKOKSKHQSQDELMADISKISQLEDE 179

QY 181 SKSLKSQIAEAKICTFKXMBERRAIAIKDALNENSQLOTHKQIFQEAEEVWKEVSE 240
DB 181 SKSLKSQIAEAKICTFKXMBERRAIAIKDALNENSQLOTHKQIFQEAEEVWKEVSE 240

QY 180 SKSLKSQIAEAKICTFKXMBERRAIAIKDALNENSQLOTHKQIFQEAEEVWKEVSE 238
DB 180 SKSLKSQIAEAKICTFKXMBERRAIAIKDALNENSQLOTHKQIFQEAEEVWKEVSE 238

QY 241 LNKOKITFEDSKVHAEVQNDENHKTGLTGLPMKQAAVLEEDTDDNLELVNSE 300
DB 241 LNKOKITFEDSKVHAEVQNDENHKTGLTGLPMKQAAVLEEDTDDNLELVNSE 300

QY 239 LNKOKITFEDSKVHAEVQNDENHKTGLTGLPMKQAAVLEEDTDDNLELVNSE 298
DB 239 LNKOKITFEDSKVHAEVQNDENHKTGLTGLPMKQAAVLEEDTDDNLELVNSE 298

QY 301 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIVQISEVDKTEELTEHKNQ 360
DB 301 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIVQISEVDKTEELTEHKNQ 360

QY 299 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIVQISEVDKTEELTEHKNQ 358
DB 299 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIVQISEVDKTEELTEHKNQ 358

QY 361 TQASLOSENTHFENENKQLOKIKWTELYQENEMKHLKTEVYENYLEKEEKLKVD 420
DB 361 TQASLOSENTHFENENKQLOKIKWTELYQENEMKHLKTEVYENYLEKEEKLKVD 420

QY 359 TQASLOSENTHFENENKQLOKIKWTELYQENEMKHLKTEVYENYLEKEEKLKVD 418
DB 359 TQASLOSENTHFENENKQLOKIKWTELYQENEMKHLKTEVYENYLEKEEKLKVD 418

QY 421 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 480
DB 421 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 480

QY 419 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 478
DB 419 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 478

QY 481 KENAHNRQKLTETELKTELLEKOPVADLVNTAFGRHSYGPSPGLGWPSSETRAFSLPP 540
DB 481 KENAHNRQKLTETELKTELLEKOPVADLVNTAFGRHSYGPSPGLGWPSSETRAFSLPP 540

QY 479 KENAHNRQKLTETELKTELLEKOPVADLVNTAFGRHSYGPSPGLGWPSSETRAFSLPP 538
DB 479 KENAHNRQKLTETELKTELLEKOPVADLVNTAFGRHSYGPSPGLGWPSSETRAFSLPP 538

QY 541 TLLEGLRLGLPLPGGGRGSRGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
DB 541 TLLEGLRLGLPLPGGGRGSRGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600

QY 539 TLLEGLRLGLPLPGGGRGSRGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 598
DB 539 TLLEGLRLGLPLPGGGRGSRGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 598

QY 601 WQDRRMWFPFGQSPDPSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 660
DB 601 WQDRRMWFPFGQSPDPSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 660

QY 599 WQDRRMWFPFGQSPDPSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 658
DB 599 WQDRRMWFPFGQSPDPSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 658

QY 661 MESSRNDTKDLGNLVNPDSSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDLGNLVNPDSSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 720

QY 659 MESSRNDTKDLGNLVNPDSSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 718
DB 659 MESSRNDTKDLGNLVNPDSSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 718

QY 721 FPPPPPGAMGASRDYFFPRFPFGPPAPFAMRNVPYPPRGPPPLPPRPPFPFPPHSEG 780
DB 721 FPPPPPGAMGASRDYFFPRFPFGPPAPFAMRNVPYPPRGPPPLPPRPPFPFPPHSEG 780

QY 719 FPPPPPGAMGASRDYFFPRFPFGPPAPFAMRNVPYPPRGPPPLPPRPPFPFPPHSEG 778
DB 719 FPPPPPGAMGASRDYFFPRFPFGPPAPFAMRNVPYPPRGPPPLPPRPPFPFPPHSEG 778

QY 781 RSEFSPGLIPSPNEPATEHPPEQQT 806
DB 779 RSEFSPGLIPSPNEPATEHPPEQQT 804

RESULT 2
MYH9_CHICK
ID MYH9_CHICK STANDARD; PRT; 1959 AA.

```

```

AC PI4105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NM_001001)
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=9004668; PubMed=2813355;
RA Shohet R.V., Corti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RA Adelstein R.S.;
RA "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
RA cellular myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989).
RL CC - FUNCTION: Cellular myosin appears to play a role in cytokinesis,
CC cell shape, and specialized functions such as secretion and
CC capping.
CC CC - SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy
CC chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2
CC regulatory light chain subunits (MLC-2).
CC CC - DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC CC - SIMILARITY: Contains 1 myosin-like globular head domain.
CC CC - SIMILARITY: Contains 1 IQ domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M26510; AAA48974.1;
DR PIR; A33977; A33977.
DR HSP; P10567; IBR2.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; Myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ_1.
DR Pfam; PF00612; Myosin_head; 1.
DR Pfam; PF00612; Myosin_N; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS50096; IQ; 1.
DR Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Multigene family; MYOSIN HEAD-LIKE.
FT DOMAIN 1 778
FT DOMAIN 779 808
FT DOMAIN 837 1925
FT NP_BIND 174 181
FT NP_BIND 654 676
FT DOMAIN 1959 AA; 226502 MW; A75C66086FD3A1A1 CRC64;
SQ SEQUENCE 1959 AA; 226502 MW; A75C66086FD3A1A1 CRC64;

Query Match 6.5%; Score 276.5; DB 1; Length 1959;
Best Local Similarity 24.5%; Pred. No. 0.0028;
Matches 133; Conservative 96; Mismatches 173; Indels 141; Gaps 26;

QY 59 LWRSEFVSRLVYVREKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELDASFE 118
DB 1299 LAKDFSALESQLODTQE-----LLQEFRLKLSFSTKLKQTE--DEKNALEQLEE 1347

```


1650 1660 1670 1680 1690 1700 1710
GGAGACCTATAGAAACCGGCAAGATCTTGAAGAGNATGGAGAGACTATTCTATCACTTATCAAGGGCA
-----ACTGACCGATATTCCTCAATAAATTAATAGGAGTGTTCCTTTGGGTTTTTTTTTGAA
660 670 680 690 700 710
1720 1730 1740 1750 1760 1770 1780
GATTATTTCCCATGCGAAGAAAGCAGCATGATTAATGGTGGCAGCTCGAATGCTGAAAGAAACCTCAATGA
AAATGT-----TGAAGAGGGAGTGTGTTTAAATTTGAATTTTTTCCAAAGCAGAGACCAAAATTTCTAG
720 730 740 750 760 770 780
1790 1800 1810 1820 1830 1840 1850
TTTAGGAAAGAAATGCTCACACAGACAAATAATTAATCT---GAAACAGAGCTTAAATTTGAATTTTGA
TTATCGGTCCGCGAATCATTTGTACCGTCAAAAGAAATTTGGTTTAAACATGATCTCTGAGTTTAAAA
790 800 810 820 830 840 850
1860 1870 1880 1890 1900 1910 1920
AAAAGATCCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCCTCCAC
ATCCGCTGAAATTTGGTTCGATTTCTCGGTTATTCA---AGCAAATGATCCGCAAGCACTGATGTTGGCG
860 870 880 890 900 910 920
1930 1940 1950 1960 1970 1980 1990
ATTGGGTTGCTTCTATCTGAAACAGAGCTTTCTCTCTCTCCAACTTTGTTGAGGGTCCACTCAGACT
AAAAGAACGAAACACGCGAAGATAGCCCTCAATGTCCAACGTGGAGATCGAGCGCTTGGTTTCGCAAGACA
930 940 950 960 970 980 990
2000 2010 2020 2030 2040 2050 2060
CTCACCTTTCTCCAGGGGGAGAGAGAGCTCACAGGCCCGAGGAATCCTCTGGACCATCAGATTAC
AGAGGAGGCTTTTAAAGCCCGCGACCGAACTGAAGAGGAGCTTCAGATGCTGAAAGCGT
1000 1010 1020 1030 1040 1050 X
2070 2080 2090 2100
CAATGAAGAGGAGATCAAGCTGTGATAGTTAACCGA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 84 Seconds
(without alignments)
3027.472 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEEPGNTPPYGLVLELR.....GLIPSNPEATBEPPEQET 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPREMBL 25: *
- 1: sp_archaea: *
- 2: sp_bacteria: *
- 3: sp_fungi: *
- 4: sp_human: *
- 5: sp_invertebrate: *
- 6: sp_mammal: *
- 7: sp_mhc: *
- 8: sp_organelle: *
- 9: sp_phage: *
- 10: sp_plant: *
- 11: sp_rodent: *
- 12: sp_virus: *
- 13: sp_vertebrate: *
- 14: sp_unclassified: *
- 15: sp_rvirus: *
- 16: sp_bacterioph: *
- 17: sp_archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3812	90.1	775	4 Q8IX93	Q8IX93 homo sapien
2	3652.5	86.3	788	4 Q86T66	Q86T66 homo sapien
3	3643	86.1	771	4 Q8IX92	Q8IX92 homo sapien
4	3626	85.7	777	4 Q95046	Q95046 homo sapien
5	3529	83.4	811	4 Q8NE23	Q8NE23 homo sapien
6	3312	78.3	732	4 Q8EUF2	Q8EUF2 homo sapien
7	3253	76.9	754	4 Q36R76	Q36R76 homo sapien
8	3235	76.5	745	4 Q8NE38	Q8NE38 homo sapien
9	2921.5	69.0	779	11 Q8R311	Q8R311 mus musculus
10	2317.5	54.8	582	11 Q8CIE3	Q8CIE3 mus musculus
11	1605	37.9	370	4 Q8IX94	Q8IX94 homo sapien
12	1542.5	36.5	365	4 Q96N90	Q96N90 homo sapien
13	1242.5	29.4	691	13 Q7ZU98	Q7ZU98 brachydanio
14	766	18.1	1193	4 Q82580	Q82580 homo sapien
15	605.5	14.3	407	13 Q7S274	Q7S274 xenopus lae
16	546	12.9	158	4 Q8IX95	Q8IX95 homo sapien

17	502	11.9	461	4 Q86Y60	Q86Y60 homo sapien
18	345	8.2	1430	5 Q9VMA7	Q9VMA7 drosophila
19	305	7.2	1326	4 Q9P219	Q9P219 homo sapien
20	286	6.8	1239	5 Q07569	Q07569 entameba h
21	285.5	6.7	1257	5 Q9VMA8	Q9VMA8 drosophila
22	276	6.5	1999	11 Q63731	Q63731 rattus norv
23	274.5	6.5	1331	4 Q86T83	Q86T83 homo sapien
24	271.5	6.4	1784	13 Q90Y85	Q90Y85 gallus gall
25	271.5	6.4	1743	5 Q96063	Q96063 dugesia jap
26	266	6.3	1427	11 Q80XQ1	Q80XQ1 mus musculus
27	266	6.3	1790	3 Q07380	Q07380 saccharomyc
28	264.5	6.3	1762	10 Q94DC2	Q94DC2 oryza sativ
29	263.5	6.2	1833	11 Q80T68	Q80T68 mus musculus
30	263	6.2	1323	4 Q8N625	Q8N625 homo sapien
31	263	6.2	1421	4 Q9P257	Q9P257 homo sapien
32	263	6.2	1636	4 Q9H357	Q9H357 homo sapien
33	262	6.2	1958	5 Q96062	Q96062 dugesia jap
34	261.5	6.2	742	5 Q813B2	Q813B2 plasmodium
35	261	6.2	1940	5 Q9U7E3	Q9U7E3 pecten maxi
36	260.5	6.2	2612	5 Q815X5	Q815X5 plasmodium
37	258.5	6.1	895	11 Q811I2	Q811I2 mus musculus
38	258.5	6.1	2007	13 Q02015	Q02015 gallus gall
39	258	6.1	2011	5 Q8MWC2	Q8MWC2 drosophila
40	258	6.1	2017	5 Q94992	Q94992 drosophila
41	258	6.1	2056	5 Q9W0W8	Q9W0W8 drosophila
42	258	6.1	2057	5 Q94987	Q94987 drosophila
43	257.5	6.1	1837	3 Q74424	Q74424 schizosacch
44	257	6.1	1831	10 Q7XN11	Q7XN11 oryza sativ
45	257	6.1	1953	3 Q13450	Q13450 saccharomyc

ALIGNMENTS

RESULT 1

Q8IX93	PRELIMINARY;	PRT;	775 AA.
ID	Q8IX93		
AC	Q8IX93;		
DT	01-MAR-2003 (TREMELrel. 23, Created)		
DT	01-MAR-2003 (TREMELrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMELrel. 23, Last annotation update)		
DE	CTAGE-5A protein		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Usener D., Schagdendorff D., Eichmuller S.;		
RT	"CTAGE-family members"		
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF338233; A007610.1.1. -		
SQ	SEQUENCE 775 AA; 87970 MW; DF5CB0BBCEA390C2 CRC64;		

Query Match	90.1%;	Score 3812;	DB 4;	Length 775;
Best Local Similarity	93.7%;	Pred. No. 1.5e-166;		
Matches 728;	Conservative 18;	Mismatches 29;	Indels 2;	Gaps 2;
QY	30	MRPDNPYGPPELVVCAAVGGFFVLLPLNRSFVSRLVYVGEQKLGATLSGLIEEK	89	
Db	1	MRPDNLXGFPWELVCAAVGGFFVLLPLNRSFVSRLVYVGEQKLGATLSGLIEEK	60	
QY	90	CKLLEKFSLIQKYEYGEVSESSLEDAPEKZAAEARSLEATCEKLNRSNSEDEILCL	149	
Db	61	SKLLEKFSLVQKYEYGEVSESSLKDAPEKAT-BAQSLKATCEKLNRSNSEDEILCL	119	
QY	150	EKDLKQEKSKHQDELMDADISKSIOSEDESKSKQIAEAKIICKTPKMSERRATAI	209	
Db	120	EKELKEKSKHQDELMDADISKRIOSEDESKSKQIAEAKMTFKIQQMEERLKIAI	179	
QY	210	KDALNENSOLOTHKOLFQCBAEVWKNKQITFEDSKVHAEQVLNDEKHITKL	269	
Db	180	KDALNENSOLOTHKOLFQCBAEVWKNKQITFEDSKVHAEQVLNDEKHITKL	238	

QY 270 TGHLPMMKQAAVLBEDTDDNDLELVNSENSENGAYLDNPPGKALKKLIHAAKLNASLK 329
Db 239 TERLLKKOKWAAMLGEDTDDNDLELVNSENSENGAYLDNPPGKALKKLIHAAKLNASLK 298
QY 330 TLEGERNQIYIQLSEVDKTKBELTEHIKULOQQAQASLOSENTHFENENKQKQKQKLVNTE 389
Db 299 TLEGERNQIYIQLSEVDKTKBELTEHIKULOQQAQASLOSENTHFENENKQKQKQKLVNTE 358
QY 390 LYQENEMKLRKLTVEENYRLEKEKLSKVDEKISHATEELETYKRAKDLESELEERTIH 449
Db 359 LYQENEMKLRKLTVEENYRLEKEKLSKVDEKISHATEELETYKRAKDLESELEERTIH 418
QY 450 SYQGGIISHEKKAHDNLAARNAERNDLRKENAHNRKLTETELKPELLEKDPYALDV 509
Db 419 SYQGGIISHEKKAHDNLAARNAERNDLRKENAHNRKLTETELKPELLEKDPYALDV 478
QY 510 PNTAFGREGSPGSPPLGWSSETRAFLSPPTLLGRLPLSLPLGCGGGRSGRSGNEID 569
Db 479 PNTAFGREGSPGSPPLGWSSETRAFLSPPTLLGRLPLSLPLGCGGGRSGRSGNEID 538
QY 570 HOITNERGESSCDRLTDHRAPSDTSGLSPWDQRRMMFPFGGSGYSDSALPPQQRDF 629
Db 539 HOITNERGESSCDRLTDHRAPSDTSGLSPWDQRRMMFPFGGSGYSDSALPPQQRDF 598
QY 630 CNSGSLGSPAELRSNMPSLDKMDGSMPESSSRNDTKDGLNLPDSSSLPAENAT 689
Db 599 CNSGSLGSPAELRSNMPSLDKMDGSMPESSSRNDTKDGLNLPDSSSLPAENAT 658
QY 690 GGFVPPPLAPIRGLFPVDAGPFLRGPFPFPFPFGMFGASRDYPPRDPFGPPAP 749
Db 659 GGFVPPPLAPIRGLFPVDAGPFLRGPFPFPFPFGMFGASRDYPPRDPFGPPAP 718
QY 750 FANRYVPPRGGPYLPFRPGFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFP 806
Db 719 FANRYVPPRGGPYLPFRPGFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFP 775

RESULT: 2
Q86TF6 PRELIMINARY; PRT; 788 AA.
AC Q86TF6;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Similar to CTAGE-2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeng H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smallegange D.E., Schnerk A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051363; AAH51363.1;
FT NON TER 1
SQ SEQUENCE 788 AA; 89147 MW; 243CA46DFD346686 CRC64;
Query Match 86.3%; Score 3652.5; DB 4; Length 788;
Best Local Similarity 93.1%; Pred. No. 3e-159;
Matches 701; Conservative 18; Mismatches 27; Indels 7; Gaps 3;
QY 59 LWSFRSVRLVYVREKQKLGATLSGLIBKCKLLEKFKSLIQKEYEGYEVESSELDASFE 118
Db 38 LQCHCLKYRSLVYVREKQKLGATLSGLIBKCKLLEKFKSLIQKEYEGYEVESSELDASFE 97
QY 119 KEAAEARSUE-----ATCEKLNRSNSELDEILCEKQKQKSKSHQODELMADISKS 173
Db 98 KEAT-EAQSLEVENQWATCEKLNRSNSELDEILCEKQKQKSKSHQODELMADISKR 156
QY 174 IQSLEDESKLSQIAEAKTICTFKMSBERRAIAIKDALNENSQLOTSKHLQFOEAEV 233
Db 157 IQSLEDESKLSQVAAEKMTFKI FQWBEELKIAIKDALNENSQLOESQKQLL-QEAEV 215
QY 234 WKSEVSELNKKITFEDSKVHAEQVLDNKHENIKITLTHLPMKQAAVLBEDTDDNDL 293
Db 216 WKEQVSELNKKITFEDSKVHAEQVLDNKHENIKITLTHLPMKQAAVLBEDTDDNDL 275
QY 294 ELEVNSENGAYLDNPPGKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKBELT 353
Db 276 ELEVNSENGAYLDNPPGKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKBELT 335
QY 354 EHKNLQQAQASLOSENTHFENENKQKQKQKLVNTELYQENEMKLRKLTVEENYRLEKE 413
Db 336 EHKNLQQAQASLOSENTHFENENKQKQKQKLVNTELYQENEMKLRKLTVEENYRLEKE 395
QY 414 EKLSKVDKISHATEELETYKRAKDLESELEERTIHSHYQGGIISHEKKAHDNLAARNAE 473
Db 396 EKLSKVDKISHATEELETYKRAKDLESELEERTIHSHYQGGIISHEKKAHDNLAARNAE 455
QY 474 RNLDLFEKNAHNRKLTETELKPELLEKDPYALDVNTAFGREGSPGSPPLGWSSET 533
Db 456 RNLDLFEKNAHNRKLTETELKPELLEKDPYALDVNTAFGREGSPGSPPLGWSSET 515
QY 534 RAFLSPPTLLGRLPLSLPLGCGGGRSGRSGNEIDTDPHRAPSD 593
Db 516 RAFLSPPTLLGRLPLSLPLGCGGGRSGRSGNEIDTDPHRAPSD 575
QY 594 TGSLSPPWDQRRMMFPFGGSGYSDSALPPQQRDFCSNSGRSLGSPAELRSFNMPSLDKM 653
Db 576 TGSLSPPWDQRRMMFPFGGSGYSDSALPPQQRDFCSNSGRSLGSPAELRSFNMPSLDKM 635
QY 654 DGSMPSESSSRNDTKDGLNLPDSSSLPAENATGPGFVPPPLAPIRGLFPVDARGP 713
Db 636 DGSMPSESSSRNDTKDGLNLPDSSSLPAENATGPGFVPPPLAPIRGLFPVDARGP 695
QY 714 FLRGGPFPFPFPFGMFGASRDYPPRDPFGPPAPFANRYVPPRGGPYLPFRPGFPFP 773
Db 696 FLRGGPFPFPFPFGMFGASRDYPPRDPFGPPAPFANRYVPPRGGPYLPFRPGFPFP 755
QY 774 PPPHSEGRSEFPFGGLIPPSNEPATEHPEPQOET 806
Db 756 PPPHSEGRSEFPFGGLIPPSNEPATEHPEPQOET 788

RESULT 3
Q8IX92
ID Q8IX92
AC Q8IX92;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 21 Seconds

(without alignments)
3691.922 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEEPGATPQYGLGLVLELR.....GLIPPSNEPATEHPPEPQQT 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286	6.8	2139	2 T18296	myosin heavy chain
2	276.5	6.5	1959	1 A33977	myosin heavy chain
3	276	6.5	1999	1 S21801	myosin heavy chain
4	274.5	6.5	1130	2 T34081	hypothetical prote
5	271.5	6.4	1875	2 S38173	myosin-like protei
6	270	6.4	1300	2 I53799	CG1 protein - huma
7	270	6.4	1356	2 S32763	kinectin 1 - huma
8	270	6.4	1410	1 A57013	early endosome ant
9	268	6.3	1957	2 T38077	hypothetical colle
10	266.5	6.3	1976	2 A59252	myosin heavy chain
11	266	6.3	1790	2 S67593	transport protein
12	258.5	6.1	2007	1 B43402	myosin heavy chain
13	258	6.1	2017	1 A36014	myosin heavy chain
14	258	6.1	2057	2 S61477	myosin II heavy ch
15	257.5	6.1	1837	2 T41023	probable nuclear p
16	257	6.1	1961	1 A61231	myosin heavy chain
17	256.5	6.1	1931	2 A52334	slow myosin heavy
18	256	6.1	841	2 A86188	hypothetical prote
19	254.5	6.0	1939	2 T18372	repeat organellar
20	254	6.0	2253	2 T30336	nuclear/mitotic ap
21	253	6.0	1938	1 A40997	myosin heavy chain
22	253	6.0	1953	2 S63244	BN11 protein - yea
23	252.5	6.0	944	2 S26710	spindlie pole body
24	252	6.0	1940	2 A59287	myosin heavy chain
25	252	6.0	2020	2 T21174	hypothetical prote
26	251	5.9	946	2 S28061	SCP1 protein - rat
27	251	5.9	1979	1 S03166	myosin heavy chain
28	251	5.9	2288	2 T23999	hypothetical prote
29	251	5.9	3225	2 I52300	giantin - human

giantin - human
interaptin - slime
kinesin-related pr
myosin heavy chain
cardiac muscle fac
myosin II heavy ch
hypothetical prote
hypothetical prote
serine/threonine-s
liver stage antige
desmoplakin I - hu
myosin heavy chain
myosin heavy chain
myosin heavy chain
nonmuscle myosin I
hypothetical prote

ALIGNMENTS

RESULT 1

T18296

myosin heavy chain - Entamoeba histolytica

C:Species: Entamoeba histolytica

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000

C:Accession: T18296

R:Guillen, N.

submitted to the EMBL Data Library, February 1997

A:Reference number: Z18865

A:Accession: T18296

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2139 <GUI>

A:Cross-references: EMBL:L03534; NID:G1850912; PID:G1850913; PIDN:AAB48065.1

C:Genetics:

A:Gene: mhca

C:Superfamily: myosin heavy chain; myosin motor domain homology

F:91-780/Domain: myosin motor domain homology <M0>

Query Match 6.8%; Score 286; DB 2; Length 2139;

Best Local Similarity 21.2%; Pred. No. 0.0068;

Matches 127; Conservative 105; Mismatches 173; Indels 194; Gaps 21;

QY	81	TLSGLTEERCKLEKFP--SLIQEYEGVEVSLESLESDASFEKEAABEARS-LEAT----	CE 133
DB	977	TLNATVKKDKITAEQESIDEKEDEITKLGDKILKEEKDDLEQDRADVATKDDIAK 1036	
QY	134	KLNRNSELE---DEILCLEKDLKQKSKHSDQDELMADISKISQLEDESKLSQIAE 190	
DB	1037	KLNKITIECEDAKDEIAKLEQLEDEENKKN---DLTNELOQTQLKLGETEKSALAAQVAA 1093	
QY	191	AKIICKTFKMSERRAIAIKDALN---ENSOLQISH-----	223
DB	1094	TK-----KASDER-----DTLSQNLNEKLTTKTKADLEKKISGLKQDYEDLED 1141	
QY	224	-----KQLPQQAENVKGEVSELNKQKITFEDESKVH 254	
DB	1142	DKNKIEGLNNAQKIKELDDEITKGADVSOYLQKQKEEVESQIAKMQEKEAIGNDVKN 1201	
QY	255	AEQVLNDKNEHIXTLTGHLPMKDDQAAVLEEDT-----TDDDNLELVNS 299	
DB	1202	KEKTIKEKELEILOSLEKL-----DETEVEKDEAKKKKEKEMKALQEKENVESKNS 1257	
QY	300	ESENGAYLDNPPPGALKKL-----IHAAKLNASLTKTLEGERNQIY-----IQ 341	
DB	1258	TERDKKLDNLDKDTQKKLDMTADNEKLKAKAKDLKLEAQLNEVDNHEKAVADAELNKK 1317	
QY	342	LSEVDKTEBELTHIKNLQTOQASLQSENTHFNENQKLOQKL-----KV 386	
DB	1318	KAQSDKELNSLKAELEALTAKKSVESKNSDKSEKAAULGEEIDQANEKLNQADLRKA 1377	
QY	387	MTELYQENEMKL-----HRKLTIV-----EEN-YRLEKEEK-LSKVDEKI 423	

```

Db 1378 TADLOEAKAEVAEQAORDKLVADNKKMTKTLEEIKARDEENTYKVENYKVLKKEADL 1437
Qy 424 SHATELETYR-----KRAKDLEBEELERTHSHYQOQIISHEK----- 460
Db 1438 ERANENWLDIEKDRMNKEQVKLSEGLKEIKDKLNAIAEKDSIFTAKKQSDADLEELN 1497
Qy 461 ----KAHDNWLAAARN-----AERNLNDLRKENAHNRQKUTETELKPFELLEKDP 504
Db 1498 KTVSEHDEVAKLNTQITKLTDRNQSABEELNELRSKADKKKKTISELEQVNELESP 1556

RESULT 2
A33977 myosin heavy chain, nonmuscle - chicken
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A33977; S06116; A43422
R:Shohet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7726-7730, 1989
A:Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular my
A:Reference number: A33977; MUID:90046668; PMID:2813355
A:Accession: A33977
A:Molecule type: mRNA
A:Residues: 1-1959 <SHO>
A:Cross-references: GB:M26510; NID:G212382; PIDN:AAA48974.1; PID:9212383
R:Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Maeaki, T.
Eur. J. Biochem. 184, 611-616, 1999
A:Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in
s.
A:Reference number: S06116; MUID:90032648; PMID:2806244
A:Accession: S06116
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 716-1008 <KAT>
A:Cross-references: GB:X17589
A:Note: this translation is not annotated in GenBank entry GGMCFMHA, release 114
R:Hodge, T.P.; Cross, R.; Kendrick-Jones, J.
J. Cell Biol. 118, 1085-1095, 1992
A:Title: Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate
A:Reference number: A43422; MUID:92381096; PMID:1512291
A:Accession: A43422
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1900-1959 <SHO>
A:Experimental source: brush border
A:Note: sequence extracted from NCBI backbone (NCBI:P111947)
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotid
F:84-764/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:552-565/Region: actin binding #status predicted
F:626-640/Region: actin binding #status predicted
F:837-1936/Domain: coiled coil #status predicted <COI>
F:837-1277/Region: S2
F:1278-1959/Region: light meromyosin
F:1937-1959/Domain: carboxyl-terminal <CET>
F:125/Modified site: N6,N6-trimethyllysine (lys) #status predicted
F:180/Binding site: ATP (lys) #status predicted
F:694,704/Active site: Cys #status predicted

Query Match 6.58; Score 276.5; DB 1; Length 1959;
Best Local Similarity 24.58; Pred. No. 0.0014;
Matches 133; Conservative 96; Mismatches 173; Indels 141; Gaps 26;

Qy 59 LWRFSRVSRLYVGRQKLGATSLGIEKCKLEKFKSLIQEYGEYVESLSLEDSAPF 118
Db 1299 LAKDFSALQSLOQTQE-----LLQETRLKLSFKSLKQTE--DEKNALKEQLEE 1347
Qy 119 KEAEERARSLEATCKENR-----SNSLEDEILCLE-----KDL----- 153
Db 1348 EEEAK--RNLEKQSVLQQQAVARKKMDGLGCLIEAEAKKXKLQKLESILQRYEEXI 1405

```

```

Qy 154 ----KOEKSK---HSQODELMADIS---KSIOLEDESKLSQTAIAKILIKTFPMSEE 203
Db 1406 AAYDKLEKTKRILQOELDDIAVDLDHQRTVSNLEKKQKKFPQLLAEEKNI--SAKYAE 1463
Qy 204 R-----RAIAIKDALNENSQLTSHKQLFQOEAEVWKGVESELNKKKIT 247
Db 1464 RDRAEAEAREKETKALSARALEEAEIQKAEELERVNKQ--FRTEMEDLMSSKDDVGKSVHE 1522
Qy 248 FEDSKVHAGVOLDNKENHIKTLTGHLPMKQDAVLEED--TTDDNLELVNSENSENGA 305
Db 1523 LEKAGALEQOQVEE-----MKTQLEEELELOATEDAKLELVNQQAAMKQ 1568
Qy 306 Y-----LDNPPKGALKLIHAAK-----LNASLKTIEGRNQIYQLSE 344
Db 1569 FDRDLGRDEONEKEKQLIRQVREMEVELEDERKORSIAVAARKKLELDLKDLESHIDT 1628
Qy 345 VDKTKSELTEHIKNLQTO-----QASIQSNTHEP-----NENKLOQKLIKWTELYQEN 394
Db 1629 ANKNRDEAIKHVKRLOAQMKDYMRLEDTRTSREILAQAENEK---KLKGM-----EA 1680
Qy 395 EM-KLHKLTVEENRYLEKEEKLKSKYDEKISH-----ATELETYRKRAKDESELER 446
Db 1681 EMIQLEELAAERAKRQAQQRDELADETANSSSGKALAMEKRRLEARIQLEEELE 1740
Qy 447 TIHSYQG--QIISHE-KKAH---DNMLAARNLNDLRKENAHNRQKUTETELKPELL 500
Db 1741 E---QGNTEIINDRLKLAQIDQNMADLNABRS--NAQKNENARQOMERQNKELKLQ 1795
Qy 501 EKD 503
Db 1796 EME 1798

RESULT 3
S21801 myosin heavy chain, neuronal [similarity] - rat
N:Alternate names: myosin II
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: S21801; S06013; S18134
R:Sun, W.; Chantler, P.D.
J. Mol. Biol. 224, 1185-1193, 1992
A:Title: Cloning of the cDNA encoding a neuronal myosin heavy chain from mammalian brain
A:Reference number: S21801; MUID:92235856; PMID:1569576
A:Accession: S21801
A:Molecule type: mRNA
A:Residues: 1-1999 <SUN>
A:Cross-references: EMBL:X62659
R:Sun, W.; Chantler, P.D.
Biochem. Biophys. Res. Commun. 175, 244-249, 1991
A:Title: A unique cellular myosin II exhibiting differential expression in the cerebral c
A:Reference number: PNO013; MUID:91151356; PMID:1998509
A:Accession: PNO013
A:Molecule type: mRNA
A:Residues: 1914-1998, 'I' <SU2>
A:Experimental source: brain
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
F:184-763/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:541-575/Region: actin binding #status predicted
F:653-675/Region: actin binding #status predicted
F:836-1999/Domain: coiled coil #status predicted <COI>
F:836-1276/Region: S2
F:1277-1999/Region: light meromyosin
F:125/Modified site: N6,N6-trimethyllysine (lys) #status predicted
F:180/Binding site: ATP (lys) #status predicted
F:693,703/Active site: Cys #status predicted
F:1916/Binding site: phosphate (Ser) (covalent) #status predicted
F:1943/Binding site: phosphate (Ser) (covalent) #status predicted

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 60 Seconds
(without alignments)
3795.555 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEEPGATPQYLGVLVLELR.....GLIPSPNEPATEHPEPQOET 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A Geneseq 29Jan04:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4231	100.0	806	3	AAY77574 Human cyt
2	3946.5	93.3	763	6	ABG74698 Human CGD
3	3897	92.1	804	6	ABG74685 Human CGD
4	3782	89.4	806	6	ABO07125 Novel hum
5	3740	88.4	881	4	ABG05280 Novel hum
6	3712	87.7	881	4	ABG20258 Novel hum
7	3705	87.6	760	6	ABO14717 Novel hum
8	3689.5	87.2	776	5	ABG4984 Human PRO
9	3689.5	87.2	776	5	ABG95590 Human ang
10	3689.5	87.2	776	5	ADD10625 Human sec
11	3689.5	87.2	776	5	ADD11585 Human sec
12	3689.5	87.2	776	5	ADD37378 Human sec
13	3689.5	87.2	776	8	ADE41586 Human sec
14	3674.5	86.8	776	4	ABG05279 Novel hum
15	3635	85.9	777	4	ABG20257 Novel hum
16	3634	85.9	777	4	AAH18372 Peptide #
17	3634	85.9	777	4	ABB37406 Peptide #
18	3634	85.9	777	4	AAH30851 Peptide #
19	3634	85.9	777	4	ABB32158 Peptide #
20	3634	85.9	777	4	ABG52211 Human liv
21	3634	85.9	777	4	ABG40185 Human pep
22	3626	85.7	777	4	AAH18368 Peptide #
23	3626	85.7	777	4	ABB37401 Peptide #
24	3626	85.7	777	4	AAH30846 Peptide #
25	3626	85.7	777	4	ABB32152 Peptide #

26	3626	85.7	777	4	ABG22697 Protein #
27	3626	85.7	777	4	AAH58083 Human bra
28	3626	85.7	777	4	ABG52207 Human liv
29	3626	85.7	777	4	AAH05968 Peptide #
30	3626	85.7	777	5	ABG40178 Human pep
31	3595	85.0	804	6	ABO07126 Novel hum
32	3593.5	84.9	1369	2	AAH24788 Human sec
33	3593.5	84.9	1369	4	AAH39009 Human sec
34	3593.5	84.9	1369	5	ABG55718 Human pol
35	3574.5	84.5	1086	4	ABG20256 Novel hum
36	3253	76.9	754	4	AAH70884 Human CTA
37	3190	75.4	798	6	ABJ19344 NOXV rela
38	3110	73.5	778	6	ABJ19345 NOXV rela
39	767	18.1	785	7	ADC37283 Nuclear f
40	766	18.1	1124	6	ABU70418 Human adi
41	766	18.1	1124	7	ADC37285 Nuclear f
42	766	18.1	1193	4	AAH25602 Human pro
43	766	18.1	1193	5	ABG61824 Prostate
44	766	18.1	1907	5	ABB82127 Human TAN
45	766	18.1	1907	7	ADE40158 Human NOV

ALIGNMENTS

RESULT 1
AAY77574
ID AAY77574 standard; protein; 806 AA.
XX

AC AAY77574;
XX
XX 08-MAY-2000 (first entry)
XX Human cytoskeletal protein (HCYT) (clone 3768043).
XX

Human cytoskeletal protein; HCYT; cell proliferation; immunological;
reproductive; nervous disorder; cytostatic; immunosuppressive; anti-HIV;
anti-diabetic; antiarteriosclerotic; dermatological; anti-inflammatory;
anti-infertility; vasotropic; cardiant.

OS Homo sapiens.
XX
XX WO200006730-A2.

XX 10-FEB-2000.
XX
XX 30-JUL-1999; 99WO-US017167.

XX 31-JUL-1998; 98US-0155185P.
XX 04-AUG-1998; 98US-0150081P.
XX 19-AUG-1998; 98US-0155228P.

XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Tang YT, Yue H, Corley NC, Guegler KJ, Azimzai Y;
XX Patterson C, Lal P, Baughn MR;

XX WPI; 2000-195297/17.
XX N-PSDB; AA258979.

XX Human cytoskeletal proteins useful for diagnosing, treating preventing
cell proliferative, immunological, reproductive, developmental and
nervous disorders.

XX Claim 1; Page 69-70; 84pp; English.

XX The invention provides human cytoskeletal proteins (HCYT) and nucleic
acids encoding the proteins. The HCYT polypeptides can be expressed by
standard recombinant methodology. The HCYT polypeptides, modulators and
antibodies are useful for treating or preventing a disorder associated
with decreased and increased expression or activity of HCYT in mammals.

XX The polypeptides are also useful for diagnosing HCYT activity disorders
such as cell proliferative, immunological, reproductive, developmental

CC and nervous disorders. Sequences AAY77569-576 represent HCYT polypeptides

XX Sequence 806 AA;

Query Match 100.0%; Score 4231; DB 3; Length 806;
Best Local Similarity 100.0%; Pred. No. 8.1e-237;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPPGATPOYLGVLVLELRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60
DB 1 MEPPGATPOYLGVLVLELRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60

QY 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGEYEVSSLEDAFPEKE 120
DB 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGEYEVSSLEDAFPEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180
DB 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180

QY 181 SKSLKSQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240
DB 181 SKSLKSQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240

QY 241 LNKQKITPDSKVHAEQVLDNENHIKLTGHLPMKQAAVLEDDTTDDNLELVNSE 300
DB 241 LNKQKITPDSKVHAEQVLDNENHIKLTGHLPMKQAAVLEDDTTDDNLELVNSE 300

QY 301 SENGAYLDNPPGALKLIIHAALKNASLKTLEGERNQIYIOLSEVDKTKBELTTHIKNLO 360
DB 301 SENGAYLDNPPGALKLIIHAALKNASLKTLEGERNQIYIOLSEVDKTKBELTTHIKNLO 360

QY 361 TQASLOSNTFHNENQKLOKQKVMTELYOENEMKLRKLTVEENYRLEKEEKLKSKVD 420
DB 361 TQASLOSNTFHNENQKLOKQKVMTELYOENEMKLRKLTVEENYRLEKEEKLKSKVD 420

QY 421 EKISHATELELYRKAKDLEELERTHSYQGOIISHEKKAHDNWLAAARNLNDLR 480
DB 421 EKISHATELELYRKAKDLEELERTHSYQGOIISHEKKAHDNWLAAARNLNDLR 480

QY 481 KENAHNRQKLTETELKFELLEKDPVALDVNTAFGREHSPYGPSPPLGWPSSETRAFLSP 540
DB 481 KENAHNRQKLTETELKFELLEKDPVALDVNTAFGREHSPYGPSPPLGWPSSETRAFLSP 540

QY 541 TLLEGRLRLSPLPGGGGRSGPGLDQITNERGSSCDRLTDPHRAPSDTGLSP 600
DB 541 TLLEGRLRLSPLPGGGGRSGPGLDQITNERGSSCDRLTDPHRAPSDTGLSP 600

QY 601 WQDRRMFPFPGQSYDPSALPPQDRFCNSGRLSGFAELRSFNMPSLDMGDMPSSE 660
DB 601 WQDRRMFPFPGQSYDPSALPPQDRFCNSGRLSGFAELRSFNMPSLDMGDMPSSE 660

QY 661 MESSRNDTKDDLGNLVDPSSLPFAENEAATGCGFVPPPLAPITGFLFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDDLGNLVDPSSLPFAENEAATGCGFVPPPLAPITGFLFPVDARGPFLRRGPP 720

QY 721 FPPPPPGMFGASRDYFPPRDPFGPPAPFAMENVYPPRGFPFPLPPRPGFPFPPHSEG 780
DB 721 FPPPPPGMFGASRDYFPPRDPFGPPAPFAMENVYPPRGFPFPLPPRPGFPFPPHSEG 780

QY 781 RSEFPFGLPPSPNEPATEHPEQOET 806
DB 781 RSEFPFGLPPSPNEPATEHPEQOET 806

RESULT 2
ABG74698
ID ABG74698 standard; protein; 763 AA.
XX
AC ABG74698;
XX
DT 10-MAY-2003 (first entry)
XX

DE Human CGDD protein 7506167CD1 SEQ ID 24.
XX Human; cell growth, differentiation and death; CGDD; cardiant; cancer;
KW cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
XX neurological disorder.
OS Homo sapiens.
XX
XX WO2003014322-A2.
XX 20-FEB-2003.
XX
XX 08-AUG-2002; 2002WO-US025465.
XX
XX 08-AUG-2001; 2001US-0311017P.
XX 17-AUG-2001; 2001US-0313070P.
XX 17-AUG-2001; 2001US-0313071P.
XX 24-AUG-2001; 2001US-0314678P.
XX 31-AUG-2001; 2001US-0316692P.
XX 07-SEP-2001; 2001US-0317913P.
XX 14-SEP-2001; 2001US-0322182P.
XX 07-DEC-2001; 2001US-0340747P.
XX 20-DEC-2001; 2001US-0342761P.
XX 29-MAR-2002; 2002US-0369129P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky MU;
PI Duggan BM, Elliott VS, Emerling JM, Forsythe IJ, Gietzen KJ;
PI Gervad AE, Grail RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
PI Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
PI Yue H;
XX
XX WPI; 2003-268197/26.
XX N-PSDB; ABQ77437.
XX
XX New polypeptide associated with cell growth, differentiation and death,
PT useful for preparing a composition for diagnosing or treating a disease
PT e.g., cardiovascular or neurological disorders of cancer.
XX
XX Claim 79; Page 236-237; 267pp; English.
XX
XX This invention describes a novel polypeptide associated with cell growth,
CC differentiation and death (CGDD) which has cardiant, cytostatic and
CC neuroprotective activity. The polypeptides of the invention are useful
CC for preparing a composition for diagnosing, or treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional CGDD e.g., cardiovascular or neurological disorders or cancer.
CC The polynucleotides of the invention can be used for gene therapy.
CC ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
CC ABQ77414-ABQ77440
XX
XX Sequence 763 AA;
SQ

Query Match 93.3%; Score 3946.5; DB 6; Length 763;
Best Local Similarity 94.3%; Pred. No. 2.3e-220;
Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;

QY 1 MEPPGATPOYLGVLVLELRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60
DB 1 MEPPGATPOYLGVLVLELRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60

QY 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGEYEVSSLEDAFPEKE 120
DB 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGEYEVSSLEDAFPEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180
DB 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180

QY 181 SKSLKSQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240
DB 181 SKSLKSQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240

Db 181 SKSLKQIAEAKICTFKMSEBERRAIAIKDALNENSLQTSKQLFQOEAEVWKGEVSE 240
 QY 241 LNKQKITFEDSKVHAEQVLDNDKENHILKTLTGHLPMKQDAVLEEDTTDDNLELVNSE 300
 Db 241 LNKQKITFEDSKVHAEQVLDNDKENHILKTLTGHLPMKQDAVLEEDTTDDNLELVNQ 300
 QY 301 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTEH:KNLQ 360
 Db 301 WENGAYLDNPPKALKKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTEH:KNLQ 360
 QY 361 TQASLQSENTHFENENKQLQKQKLYMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 420
 Db 361 TQASLQSENTHFENENKQLQKQKLYMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 420
 QY 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGOIISHEKKAHDNWLAAARNLNDLR 480
 Db 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGOIISHEKKAHDNWLAAARNLNDLR 480
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 540
 Db 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAF----- 514
 QY 541 TLLEGPLRLSPLFGGGGRSGRSGFNGPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 515 -----GRSGRSGFNGPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 QY 601 WDODRRMMFPFGQSYDPSALPQQRDRFCNSGRISGPAELBSFNMPSLDKWDGMPSE 660
 Db 558 WDODRRMMFPFGQSYDPSALPQQRDRFCNSGRISGPAELBSFNMPSLDKWDGMPSE 617
 QY 661 MESSRNDTKDDLGNLNVDPSSLSAENEAATGFGVPPPLAPIRGFLFPVDARGPFLRRGPP 720
 Db 618 MESSRNDTKDDLGNLNVDPSSLSAENEAATGFGVPPPLAPIRGFLFPVDARGPFLRRGPP 677
 QY 721 FPPPPGAMFGASRDYPPDRDPGPPPPAPFAMRNVPYPPRGFPYLPYPPRGPFPFPHSEG 780
 Db 678 FPPPPGAMFGASRDYPPDRDPGPPPPAPFAMRNVPYPPRGFPYLPYPPRGPFPFPHSEG 737
 QY 781 RSEFPSSGLIPPSNEPATEHEPEQOET 806
 Db 738 RSEFPSSGLIPPSNEPATEHEPEQOET 763

RESULT 3
 ABG74685
 ID ABG74685 standard; protein; 804 AA.
 AC ABG74685;
 AC ABG74685;
 DT 10-MAY-2003 (first entry)
 DE Human CGDD protein 6715627CD1 SEQ ID 11.
 DE Human CGDD protein 6715627CD1 SEQ ID 11.
 KW Human: cell growth, differentiation and death; CGDD; cardiact; cancer;
 KW cyrostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
 KW neurological disorder.
 OS Homo sapiens.
 PN WO2003014322-A2.
 PD 20-FEB-2003.
 XX 08-AUG-2002; 2002MO-US025465.
 XX 08-AUG-2001; 2001US-0311017P.
 PR 17-AUG-2001; 2001US-0313070P.
 PR 24-AUG-2001; 2001US-0313071P.
 PR 24-AUG-2001; 2001US-0314678P.
 PR 31-AUG-2001; 2001US-0316692P.
 PR 07-SEP-2001; 2001US-0317913P.
 PR 14-SEP-2001; 2001US-0322162P.
 PR 07-DEC-2001; 2001US-0340747P.

PR 20-DEC-2001; 2001US-0342761P.
 PR 29-MAR-2002; 2002US-0369129P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;
 PI Dugan BM, Elliott VS, Emerling BM, Forsythe IO, Gietzen KJ, Ison CH;
 PI Gorvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Reddy R;
 PI Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
 PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
 PI Yue H;
 XX WPI: 2003-268197/26.
 DR N-PSDB; ABQ77424.
 XX
 PT New polypeptide associated with cell growth, differentiation and death,
 useful for preparing a composition for diagnosing or treating a disease
 e.g., cardiovascular or neurological disorders or cancer.
 PT
 XX Claim 66; Page 213-215; 267pp; English.
 XX
 CC This invention describes a novel polypeptide associated with cell growth,
 differentiation and death (CGDD) which has cardiact, cyrostatic and
 neuroprotective activity. The polypeptides of the invention are useful
 for preparing a composition for diagnosing, or treating a disease or
 condition associated with decreased expression or overexpression of
 functional CGDD e.g., cardiovascular or neurological disorders or cancer.
 CC The polynucleotides of the invention can be used for gene therapy.
 CC ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
 CC ABQ77414-ABQ77440
 XX
 SQ Sequence 804 AA;
 SQ
 Query Match 92.1%; Score 3897; DB 6; Length 804;
 Best Local Similarity 92.1%; Fred. No. 1.8e-217; Indels 2; Gaps 2;
 Matches 747; Conservative 21; Mismatches 36;
 QY 1 MEEPQATPQYVLGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFFVVLFLW 60
 Db 1 MEEPQATPQYVLGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFFVVLFLW 60
 QY 61 RSFRSVRSRLVGRBKGLGATLSGLIEBKCLLEKFSLIQKEYEGYEVESLEDAFKE 120
 Db 61 RSFRSVRSRLVGRBKGLGATLSGLIEBKCLLEKFSLIQKEYEGYEVESLEDAFKE 120
 QY 121 AABEARSLEATCEKLNRSNLEDEILCLEKDLQKESKHQQDELMADISKSQSLDE 180
 Db 121 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLQKESKHQQDELMADISKSQSLDE 179
 QY 181 SKSLKSQIAEAKICTFKMSEBERRAIAIKDALNENSLQTSKQLFQOEAEVWKGEVSE 240
 Db 180 SKSLKSQIAEAKICTFKMSEBERRAIAIKDALNENSLQTSKQLFQOEAEVWKGEVSE 238
 QY 241 LNKQKITFEDSKVHAEQVLDNDKENHILKTLTGHLPMKQDAVLEEDTTDDNLELVNSE 300
 Db 239 LNKQKITFEDSKVHAEQVLDNDKENHILKTLTGHLPMKQDAVLEEDTTDDNLELVNSE 298
 QY 301 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTEH:KNLQ 360
 Db 299 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTEH:KNLQ 358
 QY 361 TQASLQSENTHFENENKQLQKQKLYMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 420
 Db 359 TQASLQSENTHFENENKQLQKQKLYMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 418
 QY 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGOIISHEKKAHDNWLAAARNLNDLR 480
 Db 419 EKISHATEELETYRKRAKDLSEELERTIHSYQGOIISHEKKAHDNWLAAARNLNDLR 478
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 540
 Db 479 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 538

QY 541 TLLEGPLRLSPLLPGGGGSGRGNPLDHOITNBERGSSCDRLTDPRAPSDTGLSPP 600
 Db 539 TLLEGPLRLSPLLPGGGGSGRGNPLDHOITNBERGSSCDRLTDPRAPSDTGLSPP 598
 QY 601 WQDQRMWPPPPQSQYPSDQALPQQRDFCSNGRLSGPAELRSFNMPSLDKMDGMPSE 660
 Db 599 WQDQRMWPPPPQSQYPSDQALPQQRDFCSNGRLSGPAELRSFNMPSLDKMDGMPSE 658
 QY 661 MESSRNDTKDDLGNLNVDPSSLPANEATGPGFVPPPLPIRGPLFPVDARGPFLRRGPP 720
 Db 659 MESSRNDTKDDLGNLNVDPSSLPANEATGPGFVPPPLPIRGPLFPVDARGPFLRRGPP 718
 QY 721 FPPPPGAMFGASRDYFPPDFGPPPPAPFAMRNVPYPRGFFPYLPFRPGFFPPPPHSEGG 780
 Db 719 FPPPPGAMFGASRDYFPPDFGPPPPAPFAMRNVPYPRGFFPYLPFRPGFFPPPPHSEGG 778
 QY 781 RSEFPSSGLIPPSNEPATEHPEPQOET 806
 Db 779 RSEFPSSGLIPPSNEPATEHPEPQOET 804

RESULT 4
 ABO07125
 ID ABO07125 standard; protein; 806 AA.
 XX AC ABO07125;
 XX DT 13-AUG-2003 (first entry)
 XX DE Novel human protein NOV14a.
 XX KW NOVX; autoimmune disease; allergy; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; multiple sclerosis; addiction;
 KW anxiety; pain; diabetes; glomerulonephritis; obesity;
 KW systemic lupus erythematosus; asthma; scleroderma; pancreatitis;
 KW graft versus host disease; ulcer; anaemia; cancer; trauma; infection;
 KW cardiomyopathy; atherosclerosis; hypertension; AIDS; Crohn's disease;
 KW acquired immunodeficiency syndrome; chromosomal mapping; tissue typing;
 KW forensic biology; predictive medicine; gene therapy; human.
 XX OS Homo sapiens.
 XX WO200298900-A2.
 XX PD 12-DEC-2002.
 XX PF 04-JUN-2002; 2002WO-US017558.
 XX PR 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 07-JUN-2001; 2001US-0296575P.
 PR 11-JUN-2001; 2001US-0297414P.
 PR 12-JUN-2001; 2001US-0297567P.
 PR 15-JUN-2001; 2001US-0298528P.
 PR 18-JUN-2001; 2001US-0299133P.
 PR 19-JUN-2001; 2001US-0299230P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 22-JUN-2001; 2001US-0300177P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301530P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 03-JUL-2001; 2001US-0302951P.
 PR 12-SEP-2001; 2001US-0318727P.
 PR 27-SEP-2001; 2001US-0325685P.
 PR 22-FEB-2002; 2002US-0358814P.
 PR 03-JUN-2002; 2002US-00161927.
 XX (CURA-) CURAGEN CORP.
 PA Zerhusen BD, Kekuda R, Spytek KA, Shenoy SG, Miller CE, Hjalt T;
 PI Gerlach VL, Baumgartner JC, Guo X, Gangolli EA, Vernet CAM;

PI Padigar M, Li L, Pena CEA, Gorman L, Anderson DW, Edinger SR;
 FI Patturajan M, Stone DU;
 DR N-PSDS; ACD13199.
 XX WP1; 2003-140585/13.
 PT Novel isolated NOVX polypeptide useful treating or preventing disorders
 or syndromes such as autoimmune disease, allergies, Alzheimer's disease,
 stroke, Parkinson's disease, Huntington's disease or multiple sclerosis.
 XX Claim 1; Page 128; 408pp; English.
 XX The invention describes an isolated NOVX polypeptide (I) comprising a
 sequence selected from a sequence (S1) of 1121, 635, 299, 1720, 176, 583,
 214, 395, 1098, 134, 427, 1333, 407, 806, 804, 1253, 382, 1045, 284, 496,
 506, 759, 390, 133, 215, 240, 1069, 116, 439, 1138, 477, 316, 269, 219,
 305, 406, 460, 365, 380, 829 or 326 amino acids fully defined in the
 CC specification, and the mature form of SI. (I) is useful for treating or
 CC preventing a pathology associated with (I) in a subject, preferably
 CC human, or for identifying an agent that binds to (I), where the agent is
 CC a cellular receptor or a downstream effector. (I), a polynucleotide (II)
 CC encoding (I) or an anti-(I)-antibody (V) is useful treating or preventing
 CC disorders or syndromes such as autoimmune disease, allergies, Alzheimer's
 CC disease, stroke, Parkinson's disease, Huntington's disease, multiple
 CC sclerosis, addiction, anxiety, pain, diabetes, glomerulonephritis,
 CC systemic lupus erythematosus, asthma, scleroderma, graft versus host
 CC disease, pancreatitis, obesity, ulcers, anaemia, cancer, trauma, viral,
 CC bacterial or parasitic infections, cardiomyopathy, atherosclerosis,
 CC hypertension, acquired immunodeficiency syndrome (AIDS) or Crohn's
 CC disease. (I), (II) or (V) is useful in screening assays, detection assays
 CC (e.g., chromosomal mapping, tissue typing, forensic biology), predictive
 CC medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical
 CC trials and pharmacogenomic), and in methods of treatment (e.g.,
 CC therapeutic and prophylactic). (II) is useful in gene therapy, to express
 CC (I), to detect NOVX mRNA or a genetic lesion in a NOVX gene, and to
 CC modulate NOVX activity. This is the amino acid sequence of a novel human
 CC NOV protein
 XX Sequence 806 AA;
 SQ Query Match 89.4%; Score 3782; DB 6; Length 806;
 Best Local Similarity 90.2%; Pred. No. 8.6e-211;
 Matches 727; Conservative 29; Mismatches 50; Indels 0; Gaps 0;
 QY 1 MEEPGATPQYGLVLLEELARVVAALPESMRPDPENPYGFFSELVVCVAIVGFFVLLFLW 60
 Db 1 MEEPGATPQYGLVLLEELGRVVAALPESMRPDPENPYGFFSELVVCVAIVGFFVLLFLW 60
 QY 61 RSPFSVRSRLYVGREQKLGATLSGLIEEKCKLEKFSLIQKEYEGYEVESSESDAFKE 120
 Db 61 RSPFSVRSRLYVGREQKLGATLSGLIEEKCKLEKFSLIQKEYEGYEVESSESDAFKE 120
 QY 121 AAEARSLEATCEKLNFSNSELEDEILCLEKDKQEKSKHSQOQDELMDADISKISQISLE 180
 Db 121 AAEARSLEATCEKLNFSNSELEDEILCLEKDKQEKSKHSQOQDELMDADISKISQISLE 180
 QY 181 SKLSKQIAEAKIICKTFKMSERRATAIKDALNENSQLOTSQKLFQQAAYKGVSE 240
 Db 181 SKLSKQIAEAKIICKTFKMSERRATAIKDALNENSQLOTSQKLFQQAAYKGVSE 240
 QY 241 LNKQKITFEDSKVHAEQVINDKENHIKLTGHLPMKQQAQVLEEDTDDNLELVNSE 300
 Db 241 LNKQKITFEDSKVHAEQVINDKENHIKLTGHLPMKQQAQVLEEDTDDNLELVNSE 300
 QY 301 SENGAYLDNPPKQALKKLIHAAKLNLSKLTLEGERNQIYQLSEVDKTEELTEHIKNLQ 360
 Db 301 WENGANLDDPKQALKKLIHAAKLNLSKLTLEGERNHIIQLSEVDKTEELTEHIKNLQ 360
 QY 361 TQASLOSINTHFENQKLOOKLVKMTLYQENEMKLRKLTVEENYRLEKEKLSKD 420
 Db 361 TQASLOSINTYFESQKLOOKLVKMTLYQENEMKLRKLTVEENYRLEKEKLSRVE 420
 QY 421 EKISHATELETYRKRAKDELEEDERTIHSYQGGIISHEKKAHDNWLAAARNLNDLR 480

Db 301 SENGAYLDNPPKGAUKLIIAAKLNASLKTLEGERNQIYIQLSEVDKTKTELTHIKNLQ 360
QY 361 TQASLSQSENTHFENENOKLOKQKLTWMTELYGENEMKLRKLTVEENYRLKEEKLKSKVD 420
Db 361 TQASLSQSENTHFENENOKLOKQKLTWMTELYGENEMKLRKLTVEENYRLKEEKLKSKVD 420
QY 421 EKISHATELETRYKAKDLEBELERTHSYQGOIISHEKKAHDNWLAAARNAERLNDLR 480
Db 421 EKISHATELETRYKAKDLEBELERTHSYQGOIISHEKKAHDNWLAAARNAERLNDLR 480
QY 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPPYGPSPLGWPSSSTRAFLSP 540
Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPPYGPSPLGWPSSSTRAFLSP 540
QY 541 TLLEGRLRLSPLLPGGGGSGSGNPLDQHTNTERGSSCDRLTDPHRAPSDTQSLSP 600
Db 541 TLLEGRLRLSPLLPGGGGSGSGNPLDQHTNTERGSSCDRLTDPHRAPSDTQSLSP 600
QY 601 WQDQRMFPFPPQSQSYDPSALPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQDQRMFPFPPQSQSYDPSALPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
QY 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPPGAMFGASRDYFPPRDPFPPGPPAFAMENVYPRGFPYLPFRPPFPFPPHSG 780
Db 721 FPPPPPGAMFGASRDYFPPRDPFPPGPPAFAMENVYPRGFPYLPFRPPFPFPPHSG 780
QY 781 RSEFPGLIPPSNEPATERPEPOQET 806
Db 781 RSEFPGLIPPSNEPATERPEPOQET 806

RESULT 2

US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-328-254-6

Query Match 6.4%; Score 270; DB 1; Length 2482;
Best Local Similarity 21.1%; Pred. No. 2.le-08;
Matches 124; Conservative 110; Mismatches 185; Indels 170; Gaps 22;

QY 76 QKLGATLSGLIEBKCKLEKFLSIOKEYEGY-----EVESSLEDASFEKE-AAEARSLE 129
Db 1385 QTLSSDVSELKDKTHLOEKQLSLEKDSQALSITKCELENQIAQLNKEKELLVKESELQ 1444
QY 130 ATC-----EKLNRN-----SLEDEILCLEKDKLQEKSKHSQQDELMA 168
Db 1445 ARLSSEYKLVNSKALEAALVEKGFALRLSTQEVHQLRGIEKLRVRIEADKKQL 1504
QY 169 DISKISQSLDESKLSQIAEAKIICKTFKMEERRATAIKDALNENSLOI----- 221
Db 1505 HIAEKLERERENDSLKDVENLE---RELQSEENQELVILDAENSKAEVETLKTQIEE 1561
QY 222 -----SHKQLFQOEAEVWKGVSENLKQKITF-----ED 250
Db 1562 MARSLKVPFLDLVTLRSEKENLTKIQEKQQLSGLDKLLSPKSLLEEKEQAEIQIKEE 1621
QY 251 SKVTAEQVLN-----DKE-----NHIKTLTGLHPM- 275
Db 1622 SKTAVELNQLQKELNEAVALQCGQIMKATQSLDPPIEEHQLRNSIEKLRARLEAD 1681
QY 276 MKDQAALVEEDTTDD-----DNLELEV---NSESENGAYLDNPPKGAUKLIHAA 322
Db 1682 EKKQLCVLQQLKESEHHDLLKGRVENLERELELTARTNQEHAALEAENSKEVETL--KA 1739
QY 323 KLNASLKTLEGERNQIYIQLSEVDKTKTELTHIKNLQTOQASLSQSENTHFEN----- 377
Db 1740 KIEGTQSLRG-----LELDVVTIRSEKENLTNLOKEQERISELEIINSFENILQKEQ 1795
QY 378 QKLOOKLKWMT--ELYQENEMKLRKLTV-----EEN-----YRLEKEBKLS 417
Db 1796 EKQVQKEKSSATAMEMLTQKELNERNVAALHNDQEAQKAEQNLSSQVCELELEKALLQ 1855
QY 418 KYDE-----KISHATELETRYKAKDLEBELERTHSYQGO-----IISHEKKA 462
Db 1856 GLDEAKNNYIVLQSSVNGLIQEVEDGKLEKDEEISRLKNQIQDOEQQLVSKLSQVEGE 1915
QY 463 HDNM-----LAARN-----AERNLNDLRKENAHNRQKLTETELKFELEKX 503
Db 1916 HOLWKEQNLRLNLTVLEQKIQLVQSKNASLQDTLEVLQSSYKLENE 1964

RESULT 3

US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATTNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,528
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/325,685
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/299,133
; PRIOR FILING DATE: 2001-06-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 28
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-927-28

Query Match 89.4%; Score 3782; DB 15; Length 806;
Best Local Similarity 90.2%; Pred. No. 8.5e-202;
Matches 727; Conservative 29; Mismatches 50; Indels 0; Gaps 0;

QY 1 MEEPGATQPVYGLVLELRVVAALPESMRPDENPYGFPSELVYCAAVIGFFVLLFLW 60
DB 1 MEEPGATQPVYGLVLELRVVAALPESMRPDENPYGFPSELVYCAAVIGFFVLLFLW 60
QY 61 RSFVRVRLVYGRQKLGATLSGLIECKLLEKFSLIQKEYEGYEVESSELEDAPEKE 120
DB 61 RSFVRVRLVYGRQKLGATLSGLIECKLLEKFSLIQKEYEGYEVESSELEDAPEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEHILCKDLKOEKSKHQDELMDADISKISQSLDE 180
DB 121 AAEARSLEATCEKLNRSNSELDEHILCKDLKOEKSKHQDELMDADISKISQSLDE 180
QY 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKALNENSOLOTSKOLFQCEAEVVKGEYSE 240
DB 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKALNENSOLOTSKOLFQCEAEVVKGEYSE 240
QY 241 LNKQKITPDSKVHAEQVLDNKENHIKTLTGHLPMKQAAVLEDDTTDDNLELVNSE 300
DB 241 LNKQKITPDSKVHAEQVLDNKENHIKTLTGHLPMKQAAVLEDDTTDDNLELVNSE 300
QY 301 SENGAYLDNPPGALKLIHAKKIASIKTLLEGERNQIYIQLSEVDKTKBELTHEIKNLQ 360
DB 301 WENGANDDPPGALKLIHAKKIASIKTLLEGERNQIYIQLSEVDKTKBELTHEIKNLQ 360
QY 361 TQASLQSENTHFENENQKLOKLVMTLYOENEMKLRKLTVEENYRLEKEEKLSKVD 420
DB 361 TQASLQSENTHFENENQKLOKLVMTLYOENEMKLRKLTVEENYRLEKEEKLSKVD 420
QY 421 EKISHATELETYRKRAKDLBEELERTHSYQOJISHEKKAHDNLWLAARNAERNDLR 480
DB 421 EKISHATELETYRKRAKDLBEELERTHSYQOJISHEKKAHDNLWLAARNAERNDLR 480
QY 481 KENAHNRQKITTELKFELEKDPVALDVPNTAFGRHSPPYGPSPLGWPSSETRAFLSP 540
DB 481 KENAHNRQKITTELKFELEKDPVALDVPNTAFGRHSPPYGPSPLGWPSSETRAFLSP 540
QY 541 TLLEGPLRLSPLPGGGGSGPGNPLDQHTNTERGSSCDRLTDTPHAPSDDTGLSPP 600
DB 541 TLLEGPLRLSPLPGGGGSGPGNPLDQHTNTERGSSCDRLTDTPHAPSDDTGLSPP 600
QY 601 WQDRRMFPFPGQSYPDGALPQQRDCRCSNGLSGPAELRSNMPSLDKVQDSMPSE 660
DB 601 WQDRRMFPFPGQSYPDGALPQQRDCRCSNGLSGPAELRSNMPSLDKVQDSMPSE 660
QY 661 MESSRNDTKDDI.GNLNVPOSSIPAEATGPGFVPPPLAPITGPIFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDDI.GNLNVPOSSIPAEATGPGFVPPPLAPITGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPRDRPGGPPAPFAMRNIVPBRGPPYLPBRGPPPPPHSEG 780
DB 721 FPPPPGAMFGASRDYFPRDRPGGPPAPFAMRNIVPBRGPPYLPBRGPPPPPHSEG 780

Db 721 FPPPPGAMFGASRDYFPRDRPGGPPAPFAMRNIVPBRGPPYLPBRGPPPPPHSEG 780
QY 781 RSEFPSSGLIPPSNEPATEHEPEQOET 806
Db 781 RSEFPSSGLIPPSKEPATGHPPEQOET 806

RESULT 2
US-10-081-056-336
; Sequence 336, Application US/10081056
; Publication No. US20040043927A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P235PICI
; CURRENT APPLICATION NUMBER: US/10/081,056
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/230,978
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/000,000
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/664,610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/242,922
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 09/709,238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30952
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30873
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/747,259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20

APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Azimzai, Yalda
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0568 US
CURRENT APPLICATION NUMBER: US/09/127,665
CURRENT FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 806
TYPE: PRT
ORGANISM: HOMO SAPIENS
FEATURE:
OTHER INFORMATION: 3768043, BRSTNOT24
US-09-127-665-3

Query Match 100.0%; Score 4231; DB 15; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEEPGATPQYLGVLVEELRRVVAALPESMRDPENPYGPPSELVVCAAVIGFVVLLFLW 60
Db 1 MEEPGATPQYLGVLVEELRRVVAALPESMRDPENPYGPPSELVVCAAVIGFVVLLFLW 60
Qy 61 RSPRSVRSRLYVGRQKLGATLSGLIECKKLEKFLSLIQEYEGYEVSSLEDASPEKE 120
Db 61 RSPRSVRSRLYVGRQKLGATLSGLIECKKLEKFLSLIQEYEGYEVSSLEDASPEKE 120
Qy 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDKOEKSHSQOQDELMDISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDKOEKSHSQOQDELMDISKSIOLEDE 180
Qy 181 SKLSKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKGVESE 240
Db 181 SKLSKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKGVESE 240
Qy 241 LNKQKITFEDSKVHAEQVLDNKENHKTTLTGHLPMMKQAAVLEEDTDDNLELVNSE 300
Db 241 LNKQKITFEDSKVHAEQVLDNKENHKTTLTGHLPMMKQAAVLEEDTDDNLELVNSE 300
Qy 301 SENGAYLDNPPKGALKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKGALKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Qy 361 TQASLSQSENTHFENENKQLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
Db 361 TQASLSQSENTHFENENKQLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
Qy 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
Db 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
Qy 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPYGPSPGLGWSPSSSETRAFSLSP 540
Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPYGPSPGLGWSPSSSETRAFSLSP 540
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Qy 601 WQDRRMVFPFGQSYFDSALPQQRDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQDRRMVFPFGQSYFDSALPQQRDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
Qy 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGFGVPPPLAPIRGFLFPVDARGFLARGPP 720
Db 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGFGVPPPLAPIRGFLFPVDARGFLARGPP 720

LENGTH: 806
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte Clone No: 3768043
PCT-US99-17167-6

Query Match 100.0%; Score 4231; DB 1; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEEPGATPQYLGVLVEELRRVVAALPESMRDPENPYGPPSELVVCAAVIGFVVLLFLW 60
Db 1 MEEPGATPQYLGVLVEELRRVVAALPESMRDPENPYGPPSELVVCAAVIGFVVLLFLW 60
Qy 61 RSPRSVRSRLYVGRQKLGATLSGLIECKKLEKFLSLIQEYEGYEVSSLEDASPEKE 120
Db 61 RSPRSVRSRLYVGRQKLGATLSGLIECKKLEKFLSLIQEYEGYEVSSLEDASPEKE 120
Qy 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDKOEKSHSQOQDELMDISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDKOEKSHSQOQDELMDISKSIOLEDE 180
Qy 181 SKLSKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKGVESE 240
Db 181 SKLSKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKGVESE 240
Qy 241 LNKQKITFEDSKVHAEQVLDNKENHKTTLTGHLPMMKQAAVLEEDTDDNLELVNSE 300
Db 241 LNKQKITFEDSKVHAEQVLDNKENHKTTLTGHLPMMKQAAVLEEDTDDNLELVNSE 300
Qy 301 SENGAYLDNPPKGALKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKGALKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Qy 361 TQASLSQSENTHFENENKQLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
Db 361 TQASLSQSENTHFENENKQLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
Qy 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
Db 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
Qy 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPYGPSPGLGWSPSSSETRAFSLSP 540
Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGRHSPYGPSPGLGWSPSSSETRAFSLSP 540
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Qy 601 WQDRRMVFPFGQSYFDSALPQQRDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQDRRMVFPFGQSYFDSALPQQRDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMPS 660
Qy 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGFGVPPPLAPIRGFLFPVDARGFLARGPP 720
Db 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGFGVPPPLAPIRGFLFPVDARGFLARGPP 720
Qy 721 FPPPPPGAFGASRDYFPPRDPFGPPPPAFAMENVYPPRGGFFPYLPFRPGFFPPPHSEG 780
Db 721 FPPPPPGAFGASRDYFPPRDPFGPPPPAFAMENVYPPRGGFFPYLPFRPGFFPPPHSEG 780
Qy 781 RSEFFPSGLIPPSNEPATEHPEQOET 806
Db 781 RSEFFPSGLIPPSNEPATEHPEQOET 806

RESULT 2

US-09-127-665-3

; Sequence 3, Application US/09127665

; GENERAL INFORMATION:

QY 721 FPPPPGAMGASRDYFPPDRFPGPPAPFAMENVYPPRGFPYLPBPRGFPFPPPHSEG 780
Db 721 FPPPPGAMGASRDYFPPDRFPGPPAPFAMENVYPPRGFPYLPBPRGFPFPPPHSEG 780
QY 781 RSEFPGLIPPSNEPATEHPEPQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPQET 806

RESULT 3
US-09-744-314-6
; Sequence 6, Application US/09744314
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAU, Preeti
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 PCT
; CURRENT APPLICATION NUMBER: US/09/744,314
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/127,665
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq Ver. 3.0
; SEQ ID NO 6
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 3768043
US-09-744-314-6

Query Match 100.0%; Score 4231; DB 21; Length 806;
Best Local Similarity 100.0%; Pred. No. 3 6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEPGATPQYGLVLELRVVAALPESMRDPENPYGFPSELVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLELRVVAALPESMRDPENPYGFPSELVCAAVIGFFVLLFLW 60
QY 61 RSFRSVRSRLVYVGREQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESLSDASFEKE 120
Db 61 RSFRSVRSRLVYVGREQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESLSDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDISKISQSLDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDISKISQSLDE 180
QY 181 SKLSKSQIAEAKIICKTFKMSERRAIAIKDALNENSOQTSHKQLFQOAEVWKEVSE 240
Db 181 SKLSKSQIAEAKIICKTFKMSERRAIAIKDALNENSOQTSHKQLFQOAEVWKEVSE 240
QY 241 LNKQKITPDSKVHAEQVLDNKENHIKLTGHLPMWKDQAAVLEEDTTDDNLEVNSE 300
Db 241 LNKQKITPDSKVHAEQVLDNKENHIKLTGHLPMWKDQAAVLEEDTTDDNLEVNSE 300
QY 301 SENGAYLDNPPKGAALKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKBELTHIKNLQ 360
Db 301 SENGAYLDNPPKGAALKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKBELTHIKNLQ 360
QY 361 TQASLQSENTHFENENOKLQKLVMTLEYQENMKLHRKLTVEENYRLEKEEKLKVD 420
Db 361 TQASLQSENTHFENENOKLQKLVMTLEYQENMKLHRKLTVEENYRLEKEEKLKVD 420
QY 421 EKISHATELETYRKRAKDLEELERTIHSYQGQIIISHEKKAHDNWLAAARNAERLNDLR 480

Db 421 EKISHATELETYRKRAKDLEELERTIHSYQGQIIISHEKKAHDNWLAAARNAERLNDLR 480
QY 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLOWPSETAFILSP 540
Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLOWPSETAFILSP 540
QY 541 TLLEGPIELSLPFGGGRGSGRGNPLDHOITNERGESSCDRLTDHRAPSDTGSLSP 600
Db 541 TLLEGPIELSLPFGGGRGSGRGNPLDHOITNERGESSCDRLTDHRAPSDTGSLSP 600
QY 601 WQDQRRMWFPPPGOSYDPSALPPQDRFCNSGRSLSGPAELRFSNMFSLDKMDGMPSE 660
Db 601 WQDQRRMWFPPPGOSYDPSALPPQDRFCNSGRSLSGPAELRFSNMFSLDKMDGMPSE 660
QY 661 MESSRNDTKDGLNLDNVDSSLPANENATGCGFVPPPLAPIRGELFPVDARGPILRGPP 720
Db 661 MESSRNDTKDGLNLDNVDSSLPANENATGCGFVPPPLAPIRGELFPVDARGPILRGPP 720
QY 721 PPPPPGAMFGASRDYFPPDRFPGPPAPFAMENVYPPRGFPYLPBPRGFPFPPPHSEG 780
Db 721 PPPPPGAMFGASRDYFPPDRFPGPPAPFAMENVYPPRGFPYLPBPRGFPFPPPHSEG 780
QY 781 RSEFPGLIPPSNEPATEHPEPQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPQET 806

RESULT 4
US-10-753-200-321
; Sequence 321, Application US/10753200
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/10/753,200
; CURRENT FILING DATE: 2004-01-06
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 321
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3768043CD1
US-10-753-200-321

Query Match 100.0%; Score 4231; DB 32; Length 806;
Best Local Similarity 100.0%; Pred. No. 3 6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEPGATPQYGLVLELRVVAALPESMRDPENPYGFPSELVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLELRVVAALPESMRDPENPYGFPSELVCAAVIGFFVLLFLW 60
QY 61 RSFRSVRSRLVYVGREQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESLSDASFEKE 120
Db 61 RSFRSVRSRLVYVGREQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESLSDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDISKISQSLDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDISKISQSLDE 180
QY 181 SKLSKSQIAEAKIICKTFKMSERRAIAIKDALNENSOQTSHKQLFQOAEVWKEVSE 240
Db 181 SKLSKSQIAEAKIICKTFKMSERRAIAIKDALNENSOQTSHKQLFQOAEVWKEVSE 240
QY 241 LNKQKITPDSKVHAEQVLDNKENHIKLTGHLPMWKDQAAVLEEDTTDDNLEVNSE 300

Db 241 LNKQKITFSDSKVHAEQVNDKENHKTTLTGLHPMMKQAAVLEEDTTDDNLELVNSE 300
QY 301 SENGAYLDNPPKAGALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKAGALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
QY 361 TQASLOSNTFENENQKLOQKLKVMTELYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
Db 361 TQASLOSNTFENENQKLOQKLKVMTELYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
QY 421 EKISHATELETYRKAKDLEBELERTIHSYQOGIISHEKKAHDNWLAAARNAERNDLNR 480
Db 421 EKISHATELETYRKAKDLEBELERTIHSYQOGIISHEKKAHDNWLAAARNAERNDLNR 480
QY 481 KENAHNRQKLTTELKFELEKDPYALDVNTAFGRHSFYGPSPLGWPSSETRAFLSPP 540
Db 481 KENAHNRQKLTTELKFELEKDPYALDVNTAFGRHSFYGPSPLGWPSSETRAFLSPP 540
QY 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGSSCDRLTDPHRAPSDTGSLSPP 600
Db 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGSSCDRLTDPHRAPSDTGSLSPP 600
QY 601 WQDRRMWFPFGQSYFDSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQDRRMWFPFGQSYFDSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
QY 661 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGFVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGFVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSEG 780
Db 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSEG 780
QY 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPOQET 806

RESULT 5
US-60-240-409-321
; Sequence 321, Application US/60240409
; GENERAL INFORMATION:
; APPLICANT: Buchsinder, Jenny L.
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 P
; CURRENT APPLICATION NUMBER: US/60/240,409
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 321
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 3768043CDI
US-60-240-409-321

Query Match 100.0%; Score 4231; DB 33; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENYGPFSSELVYVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENYGPFSSELVYVCAAVIGFFVLLFLW 60

QY 61 RSRFSRSLRYVREKQKLGATLSGLIECKKLEKFSLIQKEVEGEVSESSLEDASFKE 120
Db 61 RSRFSRSLRYVREKQKLGATLSGLIECKKLEKFSLIQKEVEGEVSESSLEDASFKE 120

QY 121 AAEARSLEATCEKLNRSNSEDEILCLEKDLQKSKHSQQDELWADISKSIOLEDE 180

Db 121 AAEARSLEATCEKLNRSNSEDEILCLEKDLQKSKHSQQDELWADISKSIOLEDE 180
QY 181 SKLSKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLPQOEAEVWKGVESE 240
Db 181 SKLSKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLPQOEAEVWKGVESE 240
QY 241 LNKQKITFSDSKVHAEQVNDKENHKTTLTGLHPMMKQAAVLEEDTTDDNLELVNSE 300
Db 241 LNKQKITFSDSKVHAEQVNDKENHKTTLTGLHPMMKQAAVLEEDTTDDNLELVNSE 300
QY 301 SENGAYLDNPPKAGALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKAGALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
QY 361 TQASLOSNTFENENQKLOQKLKVMTELYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
Db 361 TQASLOSNTFENENQKLOQKLKVMTELYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
QY 421 EKISHATELETYRKAKDLEBELERTIHSYQOGIISHEKKAHDNWLAAARNAERNDLNR 480
Db 421 EKISHATELETYRKAKDLEBELERTIHSYQOGIISHEKKAHDNWLAAARNAERNDLNR 480
QY 481 KENAHNRQKLTTELKFELEKDPYALDVNTAFGRHSFYGPSPLGWPSSETRAFLSPP 540
Db 481 KENAHNRQKLTTELKFELEKDPYALDVNTAFGRHSFYGPSPLGWPSSETRAFLSPP 540
QY 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGSSCDRLTDPHRAPSDTGSLSPP 600
Db 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGSSCDRLTDPHRAPSDTGSLSPP 600
QY 601 WQDRRMWFPFGQSYFDSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQDRRMWFPFGQSYFDSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
QY 661 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGFVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGFVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSEG 780
Db 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSEG 780
QY 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPOQET 806

RESULT 6
US-09-948-941-369
; Sequence 369, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO00788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-369

Query Match 93.4%; Score 3952; DB 24; Length 862;
Best Local Similarity 93.7%; Pred. No. 3.2e-216;
Matches 755; Conservative 20; Mismatches 29; Indels 2; Gaps 2;

QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENYGPFSSELVYVCAAVIGFFVLLFLW 60

Db 59 MEEFGATPQPLGLLLELRVVAALPFGMEPDNLVGFPPWELVICAUVGVFAVLLFLW 118
Qy 61 RSFRSVRSRLVYVGRQKGLGATLSGLIEBKCKLEKFLSIOKEYEGYEVESLSDASFEKE 120
Db 119 RSFRSVRSRLVYVGRQKGLGATLSGLIEBKCKLEKFLSIOKEYEGYEVESLSDASFEKE 178
Qy 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Db 179 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 237
Qy 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKEVSE 240
Db 238 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKEVSE 296
Qy 241 LNKOKITFEDSKVHAEQVLDKNDKHIKTLTGHLPMKMDQAAVLDEEDTTDDNLELVNSE 300
Db 297 LNKOKITFEDSKVHAEQVLDKNDKHIKTLTGHLPMKMDQAAVLDEEDTTDDNLELVNSE 356
Qy 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTKSELTEHIKNLQ 360
Db 357 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTKSELTEHIKNLQ 416
Qy 361 TQASLOSINTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLKEEKSVD 420
Db 417 TQASLOSINTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLKEEKSVD 476
Qy 421 EKISHATEETLYRKRAKDLEELERTIHSYQGGIISHEKKAHDNWLAAERNLNDR 480
Db 477 EKISHATEETLYRKRAKDLEELERTIHSYQGGIISHEKKAHDNWLAAERNLNDR 536
Qy 481 KENAHNRQKLTETELKELLEKDYALDVNTAFGRHSYGPSPGLWSPSETRAFLSP 540
Db 537 KENAHNRQKLTETELKELLEKDYALDVNTAFGRHSYGPSPGLWSPSETRAFLSP 596
Qy 541 TLLSGPLRLSPLPGGGGRSGRGNPLDQITNERGESSCDRLTDPHRAPSDTGSLSPP 600
Db 597 TLLSGPLRLSPLPGGGGRSGRGNPLDQITNERGESSCDRLTDPHRAPSDTGSLSPP 656
Qy 601 WDQRRMMFPFGSGYSDSALPPQDRFCNSGRSLGPAELRSFNPDLKVDGNSPSE 660
Db 657 WDQRRMMFPFGSGYSDSALPPQDRFCNSGRSLGPAELRSFNPDLKVDGNSPSE 716
Qy 661 MESSRNDTKDGLNLPDSSLPAENATGFGFVPPPLAPIRGPFVVDARGPFLRGPP 720
Db 717 MESSRNDTKDGLNLPDSSLPAENATGFGFVPPPLAPIRGPFVVDARGPFLRGPP 776
Qy 721 PPPPPGAMFGASRDYPPPPFPFPPAPFAMRVNVPFPPPPVLPFRPPPPHSEG 780
Db 777 PPPPPGAMFGASRDYPPPPFPFPPAPFAMRVNVPFPPPPVLPFRPPPPHSEG 836
Qy 781 RSEFPGLIPPSNPATEHPHPPQET 806
Db 837 RSEFPGLIPPSNPATEHPHPPQET 862

RESULT 7
PCT-US02-25465-24

; Sequence 24, Application PC/TUS0225465
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAL, Yalda
; APPLICANT: BARROSO, Ines
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRAUL, Richard C.

; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LI, Joana X.
; APPLICANT: REDDY, Roopa
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/25465
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/369,129
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7506167CD1
; PCT-US02-25465-24

Query Match 93.3%; Score 3946.5; DB 1; Length 763;
Best Local Similarity 94.3%; Pred. No. 5.6e-216;
Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;
Qy 1 MEEFGATPQPLGLLLELRVVAALPESMRDPENPYGPPSELVCAAVIGFVLLFLW 60
Db 1 MEEFGATPQPLGLLLELRVVAALPESMRDPENPYGPPSELVCAAVIGFVLLFLW 60
Qy 61 RSFRSVRSRLVYVGRQKGLGATLSGLIEBKCKLEKFLSIOKEYEGYEVESLSDASFEKE 120
Db 61 RSFRSVRSRLVYVGRQKGLGATLSGLIEBKCKLEKFLSIOKEYEGYEVESLSDASFEKE 120
Qy 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Qy 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKEVSE 240
Db 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKEVSE 240

QY 241 LNKQKITPDSKVHAEQVLDNENHKTITGHLPMKQOAAVLEDDTTDDNLELVNSE 300
 Db 241 LNKQKITPDSKVHAEQVLDNENHKTITGHLPMKQOAAVLEDDTTDDNLELVNSQ 300
 QY 301 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
 Db 301 WENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
 QY 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEEKLKSKVD 420
 Db 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEEKLKSKVD 420
 QY 421 EKISHATELETYRKRAKOLEEELERTIHSYQOIIISHEKKAHDNWLAAARNAERNLNDLR 480
 Db 421 EKISHATELETYRKRAKOLEEELERTIHSYQOIIISHEKKAHDNWLAAARNAERNLNDLR 480
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 Db 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 QY 541 TLLEGPLRLSPLPGGGGSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 541 TLLEGPLRLSPLPGGGGSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 QY 551 -----GRSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 Db 551 -----GRSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 QY 601 WDQDRMMFPFGQSYDPSALPQDRFCNSGSLGPAELRSFNMPSLDKMDGSMPS 660
 Db 558 WDQDRMMFPFGQSYDPSALPQDRFCNSGSLGPAELRSFNMPSLDKMDGSMPS 617
 QY 661 MESSRNDTKDGLNLPDSSLPAPNEATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 720
 Db 618 MESSRNDTKDGLNLPDSSLPAPNEATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 677
 QY 721 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRGFPFPPPHSE 780
 Db 678 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRGFPFPPPHSE 737
 QY 781 RSEFPGSLIPPSNEPATEHPEPQOET 806
 Db 738 RSEFPGSLIPPSNEPATEHPEPQOET 763

RESULT 8
 US-09-948-941-336
 ; Sequence 336, Application US/05948941
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/231,328
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 12618
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 336
 ; LENGTH: 804
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-948-941-336

Query Match 92.7%; Score 3923; DB 24; Length 804;
 Best Local Similarity 93.1%; Pred. No. 1.3e-214;
 Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;
 QY 1 MEEPGATPQYLGLVLEELRRVVAALPESMRDENPYGPSELVYCAAVIGFVVLFLW 60
 Db 1 MEEPGATPQYLGLVLEELRRVVAALPESMRDENPYGPSELVYCAAVIGFVVLFLW 60
 QY 61 RSRFSRSLVYRGKLGATISGLIEEKCKLEKFLSLQKEVGEYVESSLEDASFEKE 120
 Db 61 RSRFSRSLVYRGKLGATISGLIEEKCKLEKFLSLQKEVGEYVESSLEDASFEKE 120

QY 121 AAEASASLEATCEKLNRSNSELEDEILCLEKDLKQEKSKHSGQOELMADISKSIOSEDE 180
 Db 121 AT-EAOSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKHSGQOELMADISKSIOSEDE 179
 QY 181 SKSLKQIAEAKIICITFXMSBERRAIAIKDALNENSQLOTHKQLFQQAEEVWKEVSE 240
 Db 180 SKSLKQIAEAKIICITFXMSBERRAIAIKDALNENSQLOTHKQLFQQAEEVWKEVSE 238
 QY 241 LNKQKITPDSKVHAEQVLDNENHKTITGHLPMKQOAAVLEDDTTDDNLELVNSE 300
 Db 239 LNKQKITPDSKVHAEQVLDNENHKTITGHLPMKQOAAVLEDDTTDDNLELVNSE 298
 QY 301 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
 Db 299 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 358
 QY 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEEKLKSKVD 420
 Db 359 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEEKLKSKVD 418
 QY 421 EKISHATELETYRKRAKOLEEELERTIHSYQOIIISHEKKAHDNWLAAARNAERNLNDLR 480
 Db 419 EKISHATELETYRKRAKOLEEELERTIHSYQOIIISHEKKAHDNWLAAARNAERNLNDLR 478
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 Db 479 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 538
 QY 541 TLLEGPLRLSPLPGGGGSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 539 TLLEGPLRLSPLPGGGGSGSGPGLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 598
 QY 601 WDQDRMMFPFGQSYDPSALPQDRFCNSGSLGPAELRSFNMPSLDKMDGSMPS 660
 Db 599 WDQDRMMFPFGQSYDPSALPQDRFCNSGSLGPAELRSFNMPSLDKMDGSMPS 658
 QY 661 MESSRNDTKDGLNLPDSSLPAPNEATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 720
 Db 659 MESSRNDTKDGLNLPDSSLPAPNEATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 718
 QY 721 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRGFPFPPPHSE 780
 Db 719 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRGFPFPPPHSE 778
 QY 781 RSEFPGSLIPPSNEPATEHPEPQOET 806
 Db 779 RSEFPGSLIPPSNEPATEHPEPQOET 804

RESULT 9
 US-60-430-890-34
 ; Sequence 34, Application US/60490890
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Martha
 ; APPLICANT: Rupnow, Brent A.
 ; APPLICANT: Webster, Kevin R.
 ; APPLICANT: Jackson, Donald
 ; APPLICANT: Wong, Tai W.
 ; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
 ; FILE REFERENCE: D0310 PSP
 ; CURRENT APPLICATION NUMBER: US/60/490,890
 ; CURRENT FILING DATE: 2003-07-29
 ; NUMBER OF SEQ ID NOS: 2779
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34
 ; LENGTH: 804
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-430-890-34

Query Match 92.7%; Score 3923; DB 33; Length 804;
 Best Local Similarity 93.1%; Pred. No. 1.3e-214;

```
Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFVVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFVVLLFLW 60
QY 61 RSFRSVRSRLVYVGRQKGLATISGLIEEKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYVGRQKGLATISGLIEEKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKOLKQEKSHSQODELMADISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNSELDEILCLEKOLKQEKSHSQODELMADISKSIOLEDE 179
QY 181 SKLSKQIAEAKIICKTFKQSEERRAIAIKDALNENSQLOSTSHKOLFQOEAEVWKQEVSE 240
Db 180 SKLSKQIAEAKIICKTFKQSEERRAIAIKDALNENSQLOSTSHKOLFQOEAEVWKQEVSE 238
QY 241 LNKQKITFEDSKVHAEOVLNDKKNHKTILTGHLPMKQQAVALLEEDTDDNLELVNSE 300
Db 239 LNKQKITFEDSKVHAEOVLNDKKNHKTILTGHLPMKQQAVALLEEDTDDNLELVNSE 298
QY 301 SENGAYLDNPKGALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHKNLQ 360
Db 299 SENGAYLDNPKGALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHKNLQ 358
QY 361 TQASLQSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 420
Db 359 TQASLQSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 418
QY 421 EKISHATEELETYRKRAKDLLEELERTIHSYQGOIISHEKKAHNDMLAARNLNDLR 480
Db 419 EKISHATEELETYRKRAKDLLEELERTIHSYQGOIISHEKKAHNDMLAARNLNDLR 478
QY 481 KENAHNRQKLTETELKELLEKOPYALDVNTAFGRHSPYGPSPGWPSETFLSP 540
Db 479 KENAHNRQKLTETELKELLEKOPYALDVNTAFGRHSPYGPSPGWPSETFLSP 538
QY 541 TLLEGPLRLSPLPGGGGSRGSGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db 539 TLLEGPLRLSPLPGGGGSRGSGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 598
QY 601 WDQRRMMFPFGQSYFDSALPPQORQFCNSGRLSGPAELRSFNPSLDKWDGMPSE 660
Db 599 WDQRRMMFPFGQSYFDSALPPQORQFCNSGRLSGPAELRSFNPSLDKWDGMPSE 658
QY 661 MESSRNTKODLGNLVPDSSLPAENATGPGFVPPPLAIPRGFLFPVDARGPFLRRGPP 720
Db 659 MESSRNTKODLGNLVPDSSLPAENATGPGFVPPPLAIPRGFLFPVDARGPFLRRGPP 718
QY 721 FPPPPGAMFGASRDYFPPRDFGPPAPPAMRVNYPGRGPPYLPFRGPPFPFPPHSEG 780
Db 719 FPPPPGAMFGASRDYFPPRDFGPPAPPAMRVNYPGRGPPYLPFRGPPFPFPPHSEG 778
QY 781 RSEFPGLIPSPNBPATEHPPEQET 806
Db 779 RSEFPGLIPSPNBPATEHPPEQET 804

RESULT 10
PCT-US02-25465-11
; Sequence 11, Application PC/TUS0225465
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BARROSO, Ines
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROWSKI, Mark L.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
```

```
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LI, Joana X.
; APPLICANT: REDDY, Roopa
; APPLICANT: RICHARDON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/25465
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/369,129
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 6715627CD1
PCT-US02-25465-11
```

```
Query Match 92.2%; Score 3900; DB 1; Length 804;
Best Local Similarity 92.8%; Pred. No. 2.7e-213;
Matches 748; Conservative 20; Mismatches 36; Indels 2; Gaps 2;
```

```
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFVVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVVCAAVIGFVVLLFLW 60
QY 61 RSFRSVRSRLVYVGRQKGLATISGLIEEKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYVGRQKGLATISGLIEEKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKOLKQEKSHSQODELMADISKSIOLEDE 180
Db 121 AT-EAQSLEATCEKLNRSNSELDEILCLEKOLKQEKSHSQODELMADISKSIOLEDE 179
```

QY 181 SKSLKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQKFOEAQVWKEVSE 240
Db 180 SKSLKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQKFOEAQVWKEVSE 238
QY 241 LNKQKITFDSKVHAEQVNDKSHIKTLTGHLPMKQDAAVLEEDTTDDNLELVNSE 300
Db 239 LNKQKITFDSKVHAEQVNDKSHIKTLTGHLPMKQDAAVLEEDTTDDNLELVNSE 298
QY 301 SENGAYLNDPPKGAALKKLIHAAKLNASLKTLEGERNOIYIOLSEVDKTEHINKLO 360
Db 299 SENGAYLNDPPKGAALKKLIHAAKLNASLKTLEGERNOIYIOLSEVDKTEHINKLO 358
QY 361 TOQASLSQNTHPENENQKLOKQKLVMTELYGENEMKLRKLTVEENYRLEKEEKLKVD 420
Db 359 TOQASLSQNTHPENENQKLOKQKLVMTELYGENEMKLRKLTVEENYRLEKEEKLKVD 418
QY 421 EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAERNLNDR 480
Db 419 EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAERNLNDR 478
QY 481 KENAHNRQKLTETELAFELKDPYALDVNTAFGRHSYGPSPGLWSPSETRAFLSP 540
Db 479 KENAHNRQKLTETELAFELKDPYALDVNTAFGRHSYGPSPGLWSPSETRAFLSP 538
QY 541 TLLEGPLRLSPLPGGGGRSGRPGNPLDQITNERGESSCDRLTDPHRAPSDTSLSPP 600
Db 539 TLLEGPLRLSPLPGGGGRSGRPGNPLDQITNERGESSCDRLTDPHRAPSDTSLSPP 598
QY 601 WDQDRMFPFPGQSPDPSALPQRODRFCNSGRSLGPAELRSFNPGLDQWDGMPSE 660
Db 599 WDQDRMFPFPGQSPDPSALPQRODRFCNSGRSLGPAELRSFNPGLDQWDGMPSE 658
QY 661 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGVPPPLPIRGPLFPVDARGPFLRRGPP 720
Db 659 MESSRNDTKDDLGNLNVDSLSLPAENEAATGPGVPPPLPIRGPLFPVDARGPFLRRGPP 718
QY 721 FPPPPGAMFGARDVFPFPPDFGPPAPFAMENVPFRGFPPLPFRGFPFPPPHSG 780
Db 719 FPPPPGAMFGARDVFPFPPDFGPPAPFAMENVPFRGFPPLPFRGFPFPPPHSG 778
QY 781 RSEFPFGLIPSPNEPATEHPEQOET 806
Db 779 RSEFPFGLIPSPNEPATEHPEQOET 804

RESULT 11
PCT-US02-17558-28
; Sequence 28, Application PC/TUS0217558
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Miller, Charles E.
; APPLICANT: Hjal, Tord
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gangolli, Beba A.
; APPLICANT: Vernet, Corine
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Gorman, Linda
; APPLICANT: Anderson, David W.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Patturajan, Meera
; APPLICANT: Stone, David J.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-377 D (Cura 677 Other)
; CURRENT APPLICATION NUMBER: PCT/US02/17558
; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,575
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,528
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/325,685
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/299,133
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/299,230
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,177
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/318,727
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/358,814
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/301,530
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/301,550
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/302,951
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 28
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-17558-28
Query Match 89.4%; Score 3782; DB 1; Length 806;
Best Local Similarity 90.2%; Pred. No. 1.5e-206;
Matches 727; Conservative 29; Mismatches 50; Indels 0; Gaps 0;
QY 1 MEEPGATPQYLGVLBELRRVVAALPESMRPDENPYGFPSELVCAAVIGFVVLFLW 60
Db 1 MEEPGATPQYLGVLBELRRVVAALPESMRPDENPYGFPSELVCAAVIGFVVLFLW 60
QY 61 RSFRSVRSRLVYGREQKLGATLSGLIEBKCKLLEKPSLIOKEYEGYEVESLEDAFEKE 120
Db 61 RSFRSVRSRLVYGREQKLGATLSGLIEBKCKLLEKPSLIOKEYEGYEVESLEDAFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEILCEKDLKQESKHSQDELMAVDSKISQSLDE 180
Db 121 AAEARSLEATCEKLNRSNSELDEILCEKDLKQESKHSQDELMAVDSKISQSLDE 180
QY 181 SKSLKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQKFOEAQVWKEVSE 240
Db 181 SKSLKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQKFOEAQVWKEVSE 240
QY 241 LNKQKITFDSKVHAEQVNDKSHIKTLTGHLPMKQDAAVLEEDTTDDNLELVNSE 300
Db 241 LNKQKITFDSKVHAEQVNDKSHIKTLTGHLPMKQDAAVLEEDTTDDNLELVNSE 300
QY 301 SENGAYLNDPPKGAALKKLIHAAKLNASLKTLEGERNOIYIOLSEVDKTEHINKLO 360
Db 301 WENGANLDDPPKGAALKKLIHAAKLNASLKTLEGERNOIYIOLSEVDKTEHINKLO 360

601 WQDQRRMFPQSQSYPSDSALPPQDRFCNSGRSLGPAELRSFNNPSLDKMDGSMPSR 660
657 WQDQRRMFPQSQSYPSDSALPPQDRFCNSGRSLGPAELRSFNNPSLDKMDGSMPSR 716
661 MESSRNDTKDGLNLNVDPSSLPANEATGPGVPPPLAPIRGELPPVDARGPFLRRGPP 720
717 MESSRNDTKDGLNLNVDPSSLPANEATGPGVPPPLAPIRGELPPVDARGPFLRRGPP 776
721 FPPPPGAMFGASRDYFPPRDF-PGPPAPAFAMRN-VYPRGPPPYL--PPRP-GTFPP 775
777 FPPPPGAMFGASRDYFSTRGISQGGPPAPAFAMRKCLIPRGRFSPSLTFFPRAGGFPPP 836
776 PH-SEGRSEPSGLIPPSNEPATEHPEPQOS 805
837 PHFXRXEVNPPXVLIPISNEPATEHPEPQOS 867

RESULT 14
PCT-US01-08631-50617
; Sequence 50617, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hysec, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50617
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (514)..(568)
; OTHER INFORMATION: eMATRIX, accession number BL01160B, p-value=8.551e-09, raw score (

781 RGEFFPSGLIPPSKEPATGHPEPQDT 806

RESULT 13
PCT-US01-08631-35639
; Sequence 35639, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hysec, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 35639
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(881)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-35639

Query Match 88.4%; Score 3740; DB 1; Length 881;
Best Local Similarity 90.0%; Pred. No. 4,1e-204;
Matches 730; Conservative 22; Mismatches 51; Indels 8; Gaps 7;

QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFVVLFLW 60
DB 59 MEEPGATPQYGLVLEELRRVVAALPEGMRPDENPYGPPSELVCAAVIGFVVLFLW 118
QY 61 RSFRSVRSRLVYVREOKLGATLSGLIEEKCLLEKFSLIQKEYEGYEVESSELDASFEKE 120
DB 119 RSFRSVRSRLVYVREOKLGATLSGLIEEKCLLEKFSLIQKEYEGYEVESSELDASFEKE 178
QY 121 AAEEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDISKISQSLDE 180
DB 179 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDISKISQSLDE 237
QY 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLFQOEAEVWKGEVSE 240
DB 238 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLFQOEAEVWKGEVSE 296
QY 241 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQDQAAVLEEDTDDDDNLELVNSE 300
DB 297 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQDQAAVLEEDTDDDDNLELVNSE 356
QY 301 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 360
DB 357 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 416
QY 361 TQASLOSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 420
DB 417 TQASLOSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEEKLKVD 476
QY 421 EKISHATELEETYPKRAKLEELERITHSYQGOIISHEKKADNWLAAARNLNDLR 480
DB 477 EKISHATELEETYPKRAKLEELERITHSYQGOIISHEKKADNWLAAARNLNDLR 536
QY 481 KENAHNRKQKTETELKPFELLEKOPALDVNPNTAFGRHSYPGSPGLWPSSSETRAFLSPP 540
DB 537 KENAHNRKQKTETELKPFELLEKOPALDVNPNTAFGRHSYPGSPGLWPSSSETRAFLSPP 596
QY 541 TLLEGPLRLSPLIPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
DB 597 TLLEGPLRLSPLIPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 656

Db 357 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 416
Qy 361 TQASLQSENTHFENENOKLQOKLKVMTLEYQENEMKLRKLTVEENYELKEEKLKVD 420
Db 417 TEQASLQSENTHFENENOKLQOKLKVMTLEYQENEMKLRKLTVEENYELKEEKLKVD 476
Qy 421 EKISHATEELETYRKAKDLBEELEERTIHSYQOQIISHEKKAHDNWLAAARNAERNLNDLR 480
Db 477 EKISHATEELETYRKAKDLBEELEERTIHSYQOQIISHEKKAHDNWLAAARNAERNLNDLR 536
Qy 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 540
Db 537 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 596
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDQHIITNERGESSCDRLTDPHRAPSDGTGSLSP 600
Db 597 TLLEGPLRLSPLLPGGGGRSGRGNPLDQHIITNERGESSCDRLTDPHRAPSDGTGSLSP 656
Qy 601 WDQDRMMFPFPGQSPYDSDALPQQRDRFCNSGRSLGPAELRSFNMPSLDKMDGSMPS 660
Db 657 WDQDRMMFPFPGQSPYDSDALPQQRDRFCNSGRSLGPAELRSFNMPSLDKMDGSMPS 716
Qy 661 MESSRNDTKDDLGNLNVDSLSLPAENEATGPGVPPPLAPIRGPLEFVDPARGPFLRRGPP 720
Db 717 MESSRNDTKDDLGNLNVDSLSLPAENEATGPGVPPPLAPIRGPLEFVDPARGPFLRRGPP 776
Qy 721 FPPPPGAMFGASRDYPPPPRDF-FGPPPPAPFAMRN-VYPPRGPPPYL--PPRP-GFPFPP 775
Db 777 FPPPPGAMFGASRDYFESTRGISQGGPPAPFAMRKCLIPPRGFSPLTPFPRTGGFPPPP 836
Qy 776 PH-SEGSRFPSPGLIPSPNEPATEHPEPQOE 805
Db 837 PHFXRXEVNFPXVLIPTSNEPATEHPEPQOE 867

RESULT 15
US-10-236-177-180
; Sequence 180, Application US/10236177
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, John P
; APPLICANT: Anderson, David W
; APPLICANT: Berghe, Constance
; APPLICANT: Burgess, Catherine E
; APPLICANT: Catterton, Elina
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Leach, Martin D
; APPLICANT: Lepley, Denise M
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Miller, Charles E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Sheroy, Sureh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Taupier, Raymond J
; APPLICANT: Vernier, Corine AM
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-442D
; CURRENT APPLICATION NUMBER: US/10/236,177
; PRIOR FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/318,120
; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: 60/318,130
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/318,219
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/318,430
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/322,781
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,636
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,817
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/373,212
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 60/322,816
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/323,519
; PRIOR FILING DATE: 2001-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 434
; SOFTWARE: Curasequest version 0.1
; SEQ ID NO 180
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-177-180
Query Match 87.6%; Score 3705; DB 28; Length 760;
Best Local Similarity 88.7%; Pred. No. 3.3e-202;
Matches 715; Conservative 19; Mismatches 26; Indels 46; Gaps 3;
Qy 1 MEEPGATPQPVGLGLVLELRRVVAALPESMRPDENPYGFPSELVCAAVGFFVVLFLW 60
Db 1 MEEPGATPQPVGLGLLELRR----- 21
Qy 61 RSFRSVRSRLVYGRQKLGATLSGLIEKCKLLEKFSLIQKEYGYEVESLEDAFPKE 120
Db 22 -----VRSRLVYGRKLLALMLSLGLEIEKSKLEKFSLVQKEYGYEVESLKDAPFKE 76
Qy 121 AAERASLEATCEKLNRSNLEDEILCLEKDLKQKSKSQODELMADISKSQSLEDE 180
Db 77 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQKSKSQODELMADISKSQSLEDE 135
Qy 181 SKSLKSQIAEAKIICKTFMGSEERRAIAIKDALNENSQLOTSKQLFOEAFAVWKEVSE 240
Db 136 SKSLKSQVAEAKWTFKI FOMNEERLKIADALNENSQLOESQKQL-QEAEVWKEVSE 194
Qy 241 LNKQKITFEDSKVHAEQVLDNKENHKTITGHLPMKQDAVLBEDTTDDNLELVNSE 300
Db 195 LNKQKVTFEDSKVHAEQVLDNKENSHIKTITRLLKKQDAWAWLGEDITDDNLELVNSE 254
Qy 301 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 360
Db 255 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 314
Qy 361 TQASLQSENTHFENENOKLQOKLKVMTLEYQENEMKLRKLTVEENYELKEEKLKVD 420
Db 315 TEQASLQSENTHFENENOKLQOKLKVMTLEYQENEMKLRKLTVEENYELKEEKLKVD 374
Qy 421 EKISHATEELETYRKAKDLBEELEERTIHSYQOQIISHEKKAHDNWLAAARNAERNLNDLR 480
Db 375 EKISHATEELETYRKAKDLBEELEERTIHSYQOQIISHEKKAHDNWLAAARNAERNLNDLR 434
Qy 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 540
Db 435 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 494
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDQHIITNERGESSCDRLTDPHRAPSDGTGSLSP 600
Db 495 TLLEGPLRLSPLLPGGGGRSGRGNPLDQHIITNERGESSCDRLTDPHRAPSDGTGSLSP 554
Qy 601 WDQDRMMFPFPGQSPYDSDALPQQRDRFCNSGRSLGPAELRSFNMPSLDKMDGSMPS 660

Db 555 WQDRRMFPFGGOSYPDSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGMPSE 614
QY 661 MESSRNDTKDDLGNLNVDPDSSLPAENEATGPGFVPPPLAIRGFLFPVDARGPFLRRGPP 720
Db 615 MESSRNDTKDDLGNLNVDPDSSLPAENEATGPGFVPPPLAIRGFLFPVDARGPFLRRGPP 674
QY 721 FPPPPPGAMFGASRDYFPPRDPGPPAPAFAMRNVPYPPRGFPYLPDRPGFPPPPHSEG 780
Db 675 FPPPPPGAMFGASRDYFPPGDFGPPAPAFAMRNVPYPPRGFPYLPDRPGFPPPPHSEG 734
QY 781 RSEPPSGLIPPSNEPATEHPEPQOET 806
Db 735 RSEPPSGLIPPSNEPATEHPEPQOET 760

Search completed: April 20, 2004, 07:58:47
Job time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:52:48 ; Search time 18 seconds
(without alignments)

1235.532 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEFGATPQYLGVLVLELR.....GLIPSPNEPATEHPPEQOET 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 196679 seqs, 27592530 residues

Total number of hits satisfying chosen parameters: 196679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:**

- 1: /cgn2_6/ptodata/1/paa/US05 NEW COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US05 NEW COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07 NEW COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08 NEW COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09 NEW COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US10 NEW COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US60 NEW COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3946.5	93.3	763	6	US-10-486-020-24
2	3900	92.2	804	6	US-10-486-020-11
3	270	6.4	3124	1	PCT-US04-00035-99
4	256	6.1	1684	1	PCT-US04-07412-885
5	250.5	5.9	496	6	US-10-767-701-43815
6	250	5.9	2325	6	US-10-663-433-2
7	249.5	5.9	1960	7	US-60-555-303-377
8	249.5	5.9	1960	7	US-60-555-303-378
9	249.5	5.9	1960	7	US-60-555-303-379
10	249.5	5.9	1960	7	US-60-555-303-380
11	247.5	5.8	1958	6	US-10-791-666-4
12	247.5	5.8	2054	6	US-10-791-666-2
13	245	5.8	588	1	PCT-US04-02188-109
14	245	5.8	588	6	US-10-764-425-109
15	245	5.8	1652	6	US-10-796-307-578
16	245	5.8	1938	6	US-10-796-307-576
17	245	5.8	1938	6	US-10-796-307-577
18	245	5.8	1954	6	US-10-796-307-575
19	245	5.8	1972	6	US-10-796-307-579
20	245	5.8	1972	6	US-10-796-307-581
21	245	5.8	2871	7	US-60-552-390-194
22	245	5.8	2871	7	US-60-552-390-195
23	243.5	5.8	1235	6	US-10-784-004-783
24	243.5	5.8	1235	6	US-10-784-004-1111
25	243.5	5.8	1305	6	US-10-784-004-671
26	243.5	5.8	1305	6	US-10-784-004-789

ALIGNMENTS

RESULT 1

US-10-486-020-24

; Sequence 24 Application US/10486020

; GENERAL INFORMATION: Valda; BARROSO, Ines;
; APPLICANT: AZIMZAI, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BAUGHN, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann B.; GRAUL, Richard C.;
; APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; REDDY, Roopa;
; APPLICANT: RICHARDSON, Thomas W.; SPRAGUE, William W.;
; APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: YUE, Huibin

; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH

; FILE REFERENCE: PF-1126 USN

; CURRENT APPLICATION NUMBER: US/10/486,020

; PRIOR FILING DATE: 2004-02-05

; PRIOR APPLICATION NUMBER: PCT/US02/25465

; PRIOR FILING DATE: 2002-08-02

; PRIOR APPLICATION NUMBER: US 60/311,017

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/313,070

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/313,071

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/314,678

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/316,692

; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/317,913

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/322,182

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/340,747

; PRIOR FILING DATE: 2001-12-07

; PRIOR APPLICATION NUMBER: US 60/342,761

; PRIOR FILING DATE: 2001-12-20

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PERL Program

; SEQ ID NO 24

; LENGTH: 763

not
found

Sequence 1066, Ap
Sequence 1114, Ap
Sequence 754, App
Sequence 1098, Ap
Sequence 699, App
Sequence 1078, Ap
Sequence 668, App
Sequence 1064, Ap
Sequence 659, App
Sequence 1058, Ap
Sequence 662, App
Sequence 1061, Ap
Sequence 633, App
Sequence 650, App
Sequence 652, App
Sequence 757, App
Sequence 781, App
Sequence 1040, Ap
Sequence 1053, Ap

27 243.5 5.8 1305 6 US-10-784-004-1066
28 243.5 5.8 1305 6 US-10-784-004-1114
29 243.5 5.8 1515 6 US-10-784-004-754
30 243.5 5.8 1515 6 US-10-784-004-1098
31 243.5 5.8 1585 6 US-10-784-004-699
32 243.5 5.8 1585 6 US-10-784-004-1078
33 243.5 5.8 1585 6 US-10-784-004-668
34 243.5 5.8 1655 6 US-10-784-004-1084
35 243.5 5.8 1725 6 US-10-784-004-659
36 243.5 5.8 1725 6 US-10-784-004-1058
37 243.5 5.8 1795 6 US-10-784-004-662
38 243.5 5.8 1795 6 US-10-784-004-1061
39 243.5 5.8 1935 6 US-10-784-004-633
40 243.5 5.8 1935 6 US-10-784-004-650
41 243.5 5.8 1935 6 US-10-784-004-652
42 243.5 5.8 1935 6 US-10-784-004-757
43 243.5 5.8 1935 6 US-10-784-004-781
44 243.5 5.8 1935 6 US-10-784-004-1040
45 243.5 5.8 1935 6 US-10-784-004-1053

TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 7506167CD1
 US-10-486-020-24

Query Match 93.3%; Score 3946.5; DB 6; Length 763;
 Best Local Similarity 94.3%; Pred. No. 5.9e-156;
 Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;

QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60
 DB 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60

QY 61 RSFRVSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120
 DB 61 RSFRVSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMDISKISQSLDE 180
 DB 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMDISKISQSLDE 180

QY 181 SKSLKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTSKHQLFQOAEVWKGEVSE 240
 DB 181 SKSLKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTSKHQLFQOAEVWKGEVSE 240

QY 241 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQOAAVLEDTTDDNLELVNSE 300
 DB 241 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQOAAVLEDTTDDNLELVNSE 300

QY 301 SENGAYLDNPPKGAUKKIIHAAKLNASKTLLEGERNQIYIQLSEVDKTEKEITHIKNLQ 360
 DB 301 WENGAYLDNPPKGAUKKIIHAAKLNASKTLLEGERNQIYIQLSEVDKTEKEITHIKNLQ 360

QY 361 TQASLOSNTHFENENKQLOKQKLVMTELYQENEMKLRKLTVEENYRLKEEKLSKVD 420
 DB 361 TQASLOSNTHFENENKQLOKQKLVMTELYQENEMKLRKLTVEENYRLKEEKLSKVD 420

QY 421 EKISHATELETYRKAKDLEELERTIHSYQOQIISHEKKAHDNWLAAERNAERNDLR 480
 DB 421 EKISHATELETYRKAKDLEELERTIHSYQOQIISHEKKAHDNWLAAERNAERNDLR 480

QY 481 KENAHNRQKLTETELKFELEKDPVALDVPNTAFGREHSFYGPSPLGWPSPSETRAFSLPP 540
 DB 481 KENAHNRQKLTETELKFELEKDPVALDVPNTAFGREHSFYGPSPLGWPSPSETRAFSLPP 540

QY 541 TLLEGSLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 DB 515 -----GRGSRGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557

QY 601 WQDQRMFPFPQSQYPSDLSALPPQDRFCNSGRSLGPAELRSFNMPSLDKWDGSMFSE 660
 DB 558 WQDQRMFPFPQSQYPSDLSALPPQDRFCNSGRSLGPAELRSFNMPSLDKWDGSMFSE 617

QY 661 MESSRNDTKDDLGNLNVPDSSLPAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 720
 DB 618 MESSRNDTKDDLGNLNVPDSSLPAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 677

QY 721 FPPPPFCMGASRDYFPFPPDFPPPPAPFAMRNVPYPPRFPFPPYPPRFPFPPPHSEG 780
 DB 678 FPPPPFCMGASRDYFPFPPDFPPPPAPFAMRNVPYPPRFPFPPYPPRFPFPPPHSEG 737

QY 781 RSEFPSPGLIPPSNEPATEHPEPQOET 806
 DB 738 RSEFPSPGLIPPSNEPATEHPEPQOET 763

RESULT 2

US-10-486-020-11
 ; Sequence 11, Application US/10486020
 ; GENERAL INFORMATION:
 ; APPLICANT: AZIMZAI, Yalda; BARROSO, Ines;

APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;
 APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
 APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
 APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
 APPLICANT: GORVAD, Ann E.; GRAUL, Richard C.;
 APPLICANT: GRIPEIN, Jennifer A.; GURURAJAN, Rajagopal;
 APPLICANT: KAFALIA, April J.A.; ISON, Craig H.;
 APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
 APPLICANT: LEE, Sally; LEE, Soo Yeun;
 APPLICANT: LI, Joana X.; REDDY, Roopa;
 APPLICANT: RICHARDSON, Thomas W.; SPRAGUE, William W.;
 APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
 APPLICANT: WARREN, Bridget A.; XU, Yuming;
 APPLICANT: YAO, Monique G.; YUE, Henry;
 APPLICANT: YUE, Huidun

TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
 FILE REFERENCE: PF-1126 USN
 CURRENT APPLICATION NUMBER: US/10/486,020
 CURRENT FILING DATE: 2004-02-05

PRIOR APPLICATION NUMBER: PCT/US02/25465
 PRIOR FILING DATE: 2002-08-02
 PRIOR APPLICATION NUMBER: US 60/311,017
 PRIOR FILING DATE: 2001-08-08
 PRIOR APPLICATION NUMBER: US 60/313,070
 PRIOR FILING DATE: 2001-08-17
 PRIOR APPLICATION NUMBER: US 60/313,071
 PRIOR FILING DATE: 2001-08-17
 PRIOR APPLICATION NUMBER: US 60/314,678
 PRIOR FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: US 60/316,692
 PRIOR FILING DATE: 2001-08-31
 PRIOR APPLICATION NUMBER: US 60/317,913
 PRIOR FILING DATE: 2001-09-07
 PRIOR APPLICATION NUMBER: US 60/322,182
 PRIOR FILING DATE: 2001-09-14
 PRIOR APPLICATION NUMBER: US 60/340,747
 PRIOR FILING DATE: 2001-12-07
 PRIOR APPLICATION NUMBER: US 60/342,761
 PRIOR FILING DATE: 2001-12-20

Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 54
 SOFTWARE: PERL Program
 SEQ ID NO 11
 LENGTH: 804
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 6715627CD1
 US-10-486-020-11

Query Match 92.8%; Score 3900; DB 6; Length 804;
 Best Local Similarity 92.8%; Pred. No. 5.5e-156;
 Matches 748; Conservative 20; Mismatches 36; Indels 2; Gaps 2;

QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60
 DB 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60

QY 61 RSFRVSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120
 DB 61 RSFRVSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMDISKISQSLDE 180
 DB 121 AT-EAQSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMDISKISQSLDE 179

QY 181 SKSLKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTSKHQLFQOAEVWKGEVSE 240
 DB 180 SKSLKSQIAEAKIICTFKMSEERRAIAIKDALNENSQLOTSKHQLFQOAEVWKGEVSE 238

QY 241 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQOAAVLEDTTDDNLELVNSE 300
 DB 241 LNKQKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQOAAVLEDTTDDNLELVNSE 300

Db 239 LNKQKVTFFEDSKVHAEQVINDKESHKILTERLLKXKDWAAAMLGEDITDDDDNLEMMNSE 298
Qy 301 SENGAYLDNPPGALKKLIHAAKLNASKLTGEBNQIYIOLSEVDTKTELTETIHKNLQ 360
Db 299 SENGAYLDNPPGALKKLIHAAKLNASKLTGEBNQIYIOLSEVDTKTELTETIHKNLQ 358
Qy 361 TQOASLOSTHFNENQKLOKLVMTLYOENEMKLRKLTVEENYRLEKEEKLKSV 420
Db 359 TQOASLOSTHFNENQKLOKLVMTLYOENEMKLRKLTVEENYRLEKEEKLKSV 418
Qy 421 EKISHATELETYRKAKDLEBELERTHSYOGIISHEKKAHDNWLAAARNAERLNDLR 480
Db 419 EKISHATELETYRKAKDLEBELERTHSYOGIISHEKKAHDNWLAAARNAERLNDLR 478
Qy 481 KENAHNRQKLTETELKFELLEDKPYALDVPNTAFGRHSYGPSPGLWPSSETRAFLSP 540
Db 479 KENAHNRQKLTETELKFELLEDKPYALDVPNTAFGRHSYGPSPGLWPSSETRAFLSP 538
Qy 541 TLLEGRLRLSPILPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHAPSDTGLSP 600
Db 539 TLLEGRLRLSPILPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHAPSDTGLSP 598
Qy 601 WDQDRRMFPPPPQSYDPSALPQDRFCNSGRSLSGPAELRSFNMPSLDXDGMSPSE 660
Db 599 WDQDRRMFPPPPQSYDPSALPQDRFCNSGRSLSGPAELRSFNMPSLDXDGMSPSE 658
Qy 661 MESSRNDTKDDGLNLYNPSSLPAAENATGPGVPPPLAPIRGIPPLVDARGPFLRGPP 720
Db 659 MESSRNDTKDDGLNLYNPSSLPAAENATGPGVPPPLAPIRGIPPLVDARGPFLRGPP 718
Qy 721 FPPPPPGAMFGASRDYFPDPDPGPPAPFAMNVYPPRGFPYLPFRGFPFPPHSG 780
Db 719 FPPPPPGAMFGASRDYFPDPDPGPPAPFAMNVYPPRGFPYLPFRGFPFPPHSG 778
Qy 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 779 RSEFPGLIPPSNEPATEHPEPOQET 804

RESULT 3
PCT-US04-00035-99
; Sequence 99, Application PC/TUS0400035
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; FILE REFERENCE: AM100927 (031836-002000)
; CURRENT APPLICATION NUMBER: PCT/US04/00035
; CURRENT FILING DATE: 2004-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 99
; LENGTH: 3114
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-00035-99

Query Match 6.4%; Score 270; DB 1; Length 3114;
Best Local Similarity 21.1%; Pred. No. 0.00062;
Matches 124; Conservative 110; Mismatches 185; Indels 170; Gaps 22;

Qy 76 QKLGATSLGIEEKKLEKFLKPSLIQKEVEGY-----EVESSELDASPEKE-AAEARSLE 129
Db 2017 QTLSSDVSELKDKTHLQKLSLEKDSQALSUTKCELENQIAQLNKEKELLVKESESUQ 2076
Qy 130 ATC-----EKLNRN-----SELEIDLCLEKDLKOSKSHSQDELVA 168
Db 2077 ARLSSEYKLVNSKALEAALVEKGEFALRLSSQEEVHQLRGIEKLRVRIEADKKQL 2136

Qy 169 DISKSIQSLDESKSLKQIAEAKIICKTFKMSERRAIRAKDALNENSOLQF----- 221
Db 2137 HIAEKLKERENDSLKDKVENLE---RELQMSSENOELVILDAENSKAEVETLKTQIEE 2193
Qy 222 -----SHKOLFQOEAEVVKGEVSELNKKQKITF-----ED 250
Db 2194 MARSLKVFEVDLVTLRSEKENLTQIOEQGQSLDKLLSSFKSLLEKEQEAEIQIEE 2253
Qy 251 SKVHAEQVNLN-----DKE-----NHIKTLTGHLP- 275
Db 2254 SKTAVEMLQNLKELNEVAALCGDQELMKATQSLDPPTEEBEHQLRNSIEKLRARLEAD 2313
Qy 276 MKQAAVLEBDTDD-----DNLELEV---NSENGAYLDNPPGALKKLIHAA 322
Db 2314 EKQOLCVLOQLKESEHRAHLKGRVENLERELIARTNQHAALAEANSKGEVETL--KA 2371
Qy 323 KINASLTKTGERNQIYIOLSEVDTKTELTETIHKNLQOQASLOSTHFN-----EN 377
Db 2372 KIEGWTQSLRG-----LELDVVVTIRSEKENLTNLOKEQERISELEINSSFENILQKEQ 2427
Qy 378 QKLOKLVMT--ELYOENEMKLRKLTV-----EEN-----YRLEKEEKLK 417
Db 2428 EKQVMEKESSTAMEMLQTLKELNERNVAALHNDQEAACKAEQNLSSQVECLEKEAQLLQ 2487
Qy 418 KVDE-----KISHATELETYRKAKDLEEELEERTIHSYQO-----IISHEKKA 462
Db 2488 GLDEAKNYIYVLOSSVNGLIQEVEDGKQLEKKDEEISRLKQIQDOEQOLVSKLSQVEGE 2547
Qy 463 HDNW-----LAARN-----ABRNLDLRKENAHNRQKLTETELKFELLEDK 503
Db 2548 HQLWKEQNLRLNLTVELEQKIQVLSQKNASLQDTLEVLOSSEYKNLENE 2596

RESULT 4
PCT-US04-07412-885
; Sequence 885, Application PC/TUS0407412
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Zhang Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Goodrich, Kyle W.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Gezhi
; APPLICANT: Wang, Zhiwei
; APPLICANT: Boyle, Bryan J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 822CIP/PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07412
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 10/389,559
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/365,264
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/340,187
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 10/296,115
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:44:16 ; Search time 7714 Seconds
(without alignments)
12248.372 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggtttaatccatgaagaag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:**

2: em_esthum:**

3: em_estinu:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_estl:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_phg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2654	83.9	2972	11 BC039017	BC039017 Homo sapi
2	2191.2	69.3	2666	11 BC030655	BC030655 Homo sapi
3	2001.8	63.3	2415	29 AY409371	AY409371 Homo sapi
4	1694	53.5	2415	29 AY409372	AY409372 Pan trogl

5	1611	50.9	2784	11 AK029466	AK029466 Mus muscu
6	1357.6	42.9	2334	29 AY409373	AY409373 Mus muscu
7	1352.6	42.7	1955	11 BC018567	BC018567 Mus muscu
8	1061.8	33.6	1241	11 BC029513	BC029513 Homo sapi
9	990	31.3	1067	13 BX328783	BX328783 BX328783
10	916.6	29.0	966	13 BX329133	BX329133 BX329133
11	913.6	28.9	1089	12 BM563897	BM563897 AGENCOURT
12	857.6	27.1	882	13 BX376997	BX376997 BX376997
13	800.6	25.3	936	13 BX328066	BX328066 BX328066
14	793.8	25.1	1071	13 BX400140	BX400140 BX400140
15	781.8	24.7	882	13 BX376998	BX376998 BX376998
16	737.8	23.3	785	14 CB852737	CB852737 UI-CF-FNO
17	733.4	23.2	1201	13 BX377913	BX377913 BX377913
18	732.8	23.2	736	13 BX111137	BX111137 BX111137
19	731	23.1	744	12 BG719803	BG719803 602691209
20	727.6	23.0	885	13 BX326021	BX326021 BX326021
21	725.8	22.9	918	12 BG289351	BG289351 602387471
22	724	22.9	780	12 BG740463	BG740463 602633857
23	721.2	22.8	744	13 BU687930	BU687930 UI-CF-EC1
24	716	22.6	975	13 BX391394	BX391394 BX391394
25	714.8	22.6	866	13 BX30661	BX30661 BX30661
26	712.4	22.5	822	9 AUI35481	AUI35481 AUI35481
27	703	22.2	755	13 BU942235	BU942235 AGENCOURT
28	689.4	21.8	842	12 BI561045	BI561045 603254248
29	675	21.3	912	10 BE781656	BE781656 601470531
30	668.2	21.1	687	12 BI461930	BI461930 603204268
31	664.6	21.0	752	12 BI462084	BI462084 603205263
32	649	20.5	926	10 BF980666	BF980666 602304385
33	648.6	20.5	847	13 BX368123	BX368123 BX368123
34	647.8	20.5	1178	11 AK053010	AK053010 Mus muscu
35	647.4	20.5	973	12 BG117369	BG117369 603346374
36	645.4	20.4	812	13 BU928057	BU928057 AGENCOURT
37	640.2	20.2	931	13 BQ278480	BQ278480 AGENCOURT
38	639.6	20.2	646	12 BI829591	BI829591 603079328
39	630.4	19.9	1059	13 BX367676	BX367676 BX367676
40	625.4	19.8	628	9 A1082250	A1082250 ox79h01.x
41	624.8	19.7	786	12 BI761502	BI761502 603046640
42	624.4	19.7	648	14 CA414131	CA414131 UI-H-EZ0
43	624.4	19.7	708	12 BM977364	BM977364 UI-CF-EN1
44	622.6	19.7	849	10 BF794269	BF794269 602255469
45	622.4	19.7	687	9 AA931824	AA931824 om83a04.s

ALIGNMENTS

BC039017 2972 bp mRNA linear HTC 19-NOV-2003
Homo sapiens meningioma expressed antigen 6 (coiled-coil
proline-rich), mRNA (cDNA clone IMAGE:5770231), containing
frame-shift errors.
BC039017 GI:24659234
HTC.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2972)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,B.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

TITLE	Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, O., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	22388257
REFERENCE	12477932
AUTHORS	2 (bases 1 to 2972)
JOURNAL	Straussberg, R.
REMARK	Direct Submission
COMMENT	Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
	NIH-MGC Project URL: http://mgc.ncl.nih.gov
	Contact: MGC Help desk
	Email: cgabp3@mail.nih.gov
	Tissue Procurement: Life Technologies, Inc.
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
	Web site: http://www.nisc.nih.gov/
	Contact: nisc.mgc@nih.gov
	Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granitz, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masidillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantrisp, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
source	Series: IRAX Plate: 82, Row: 1 Column: 9
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5174560
	This clone has the following problem: frame shifted.
	Location/Qualifiers
	1..2972
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clones="IMAGE:5770231"
	/tissue_type="Brain, fetal, whole pooled"
	/clone_lib="NIH MGC_121"
	/lab_host="DH10B"
	/note="vector: pCMV-SPORT6"
ORIGIN	
Query Match	83.9%; Score 2654; DB 11; Length 2972;
Best Local Similarity	95.6%; Pred. No. 0;
Matches 2765; Conservative	0; Mismatches 84; Indels 42; Gaps 3;
QY	310 GGCCAGGTACTCGGTGACCGGACAGAGAGCTCCACGCTATGGAGGAGCCCTGGTGTCT
Db	52 GGCCGGGGTACTGTGGCCACACAGAGAGCTTTGGCGCTATGGAGAGCCCGGGCT
QY	370 ACCCTCAGCCCTACCTGGGGCTGGTCTGTGAGAGAGCTACGAGAGTTGTGGCAGCACTA
Db	112 ACCCTCAACCGTATTGGGGCTGCTCTGAGAGAGCTACGAGGGTGTGGCAGCACTG
QY	430 CCTGAGCTATGAGACGAGATGAGATCTTATCGTTTCCATCGGAACCTGGTGGTATGT
Db	172 CCTGAGCTATGAGACGAGATCTTAACTTTATCGTTTCCATCGGAATTTGGTGATATGT
QY	490 GCAGCTGTATTGGATTTTGTGTTCTCTCTTTTGTGGAGAGTTTATGATCGGTT
Db	291
QY	550 AGGAGTCGGCTTTTACCTGGGAGAGAGCAAAAACCTTTGGTGCACACGCTTTCTGGACTAAT
Db	292 AGGAGTCGGCTTTTATGTGGGAGAGAGAGAGAGAGCTTTGCTCTAATGCTTTCTGGACTAAT
QY	610 GAAGAAAAATGTAACTACTTTGAAAAATTTAGCTTTATTTCAAAAAGAGTATGAGGCTAT
Db	352 GAAGAAAAATGTAACTACTTTGAAAAATTTAGCTTTATTTCAAAAAGAGTATGAGGCTAT
QY	670 GAAGTAGAGTCATCTTTTAGAGGATGCGCAGCTTTGAGAAGGAGGAGCAGAGAACGCA
Db	412 GAAGTAGAGTCATCTTTTAGAGGATGCGCAGCTTTGAGAAGGAGGCAAC---AGAAGCACA
QY	730 AGTTGGAGCAACCTGTGAAAAAGCTGAACAGGTCCTCAATTCGAACTTGAAGGATGAATC
Db	469 AGTTGGAGCAACCTGTGAAAAAGCTGAACAGGTCCTCAATTCGAACTTGAAGGATGAATC
QY	790 CTCTGCTAGAAAAAGACTTTAAAAACAAGAGAAATCTAAACATTTCTCAACAAGATGAATG
Db	529 CTCTGCTAGAAAAAGACTTTAAAAACAAGAGAAATCTCAACATTTCTGAACAAGATGAATG
QY	850 ATGCGGATATTTCAAAAAGATATACAGTCTCTTAGAAGATGAGTCAAAATCCCTCAATCA
Db	589 ATGCGGATATTTCAAAAAGATATACAGTCTCTTAGAAGATGAGTCAAAATCCCTCAATCA
QY	910 CAATATGCTGAGCCCAAAATCATCTGCAAGACATTTTAAAAATGAGTGAAGAACGACGGCT
Db	649 CAATATGCTGAGCCCAAAATCATCTGCAAGATTTTCAATGATGATGAGAACGCTGAAG
QY	970 ATACCAATAAAGATGCTTTTGAATGAAAAATCTCAACTTCAGCAAGCCCAATCAAGCTT
Db	709 ATACCAATAAAGATGCTTTTGAATGAAAAATCTCAACTTCAGCAAGCCCAATCAAGCTT
QY	1030 TTTGAGCAAGAGCTGAGTATGAGAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
Db	769 TTTGAGCAAGAGCTGAGTATGAGAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
QY	1080 ACATTTGAGCAACTCAAAAGTACACGCAAGACAAAGTCTGATGATGATGATGATGATGATG
Db	826 ACATTTGAGCAACTCAAAAGTACACGCAAGACAAAGTCTGATGATGATGATGATGATGATG
QY	1150 AAGACCTGAGTGCACACTTCCCAATGATGAGAGATCAGCTGCTGCTGCTGCTGCTGCTGCTG
Db	886 AAGACCTGAGTGCACACTTCCCAATGATGAGAGATCAGCTGCTGCTGCTGCTGCTGCTGCTG
QY	1210 ACAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db	946 ATACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
QY	1270 TTA-----
Db	1006 TTAGGTATTAAGTCACTCATCTCTCTTTTGTGCTTAAATGATAATCTCTCAAAAGGAGCT
QY	1294 TTGAAGAACTGATTCATGCTGCTAAAGTAAATGCTTTTAAATAACCTTTAGAAAGGAA
Db	1086 TTGAAGAACTGATTCATGCTGCTAAAGTAAATGCTTTTAAATAACCTTTAGAAAGGAA
QY	1354 AGAAACCAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT
Db	1126 AGAAACCAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT
QY	1414 CATATTTAAATTTCTTCAGACTCAACAGCATCTTTTGGAGTCAGAAAAACACACATTTTGA
Db	1186 CATATTTAAATTTCTTCAGACTCAACAGCATCTTTTGGAGTCAGAAAAACACACATTTTGA
QY	1474 AATGAGATCAGAGCTTCAACAGAACTTAAAGTATGATGATGATGATGATGATGATGATGATG
Db	1246 AATGAGATCAGAGCTTCAACAGAACTTAAAGTATGATGATGATGATGATGATGATGATGATG
QY	1534 GAATGAACTCCACAGGAAATTAACAGTGAAGGAAATTTATCGTTAGAGAAAAAGAG
Db	1306 GAATGAACTCCACAGGAAATTAACAGTGAAGGAAATTTATCGTTAGAGAAAAAGAG

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 01:51:06 ; Search time 12060 Seconds
(without alignments)
11371.257 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggtttaaactgaagaag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl :

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sta.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2832	89.5	59130	9	AL135904	AL135904 Human DNA
C 2	2832	89.5	130030	9	AC004889	AC004889 Homo sapi
C 3	2832	89.5	140915	9	AC005587	AC005587 Homo sapi
C 4	2830.4	89.5	130030	9	AC004889	AC004889 Homo sapi
C 5	2804.2	88.6	157801	2	AC144412	AC144412 Homo sapi
C 6	2804.2	88.6	169723	9	AC106862	AC106862 Homo sapi
C 7	2804.2	88.6	175228	2	AC143343	AC143343 Homo sapi
C 8	2804.2	88.6	200123	9	AC073264	AC073264 Homo sapi
C 9	2764.2	87.4	3676	9	HSU34780	U94780 Human menin
C 10	2749.8	86.9	162320	2	AC076959	AC076959 Homo sapi
C 11	2603.6	82.3	202634	2	AC147075	AC147075 Pan trogl
C 12	2593.8	82.0	2931	9	HSM807156	EX640994 Homo sapi
C 13	2583	81.6	2957	9	HSM808160	EX648014 Homo sapi
C 14	2531.8	80.0	2890	9	HSM803540	AL832233 Homo sapi
C 15	2531.4	80.0	2895	9	BC064355	BC064355 Homo sapi
C 16	2503.8	79.1	158310	9	AL162571	AL162571 Human DNA
C 17	2496.4	78.9	150942	9	AC017070	AC017070 Homo sapi
C 18	2490.2	77.1	199636	2	AP001126	AL139123 Human DNA
C 19	2439.4	77.1	200606	9	AC099849	AC099849 Homo sapi
C 20	2439.4	76.6	178151	9	AL136123	AL136123 Human DNA
C 21	2422.2	76.6	349980	6	AX711880	AX711880 Sequence
C 22	2422.2	76.6	161790	2	AC022371	AC022371 Homo sapi
C 23	2415	76.3	161863	9	AL606970	AL606970 Human DNA
C 24	2415	76.3	2814	6	AX335827	AX335827 Sequence
C 25	2414.6	76.3	2814	6	AX410773	AX410773 Sequence
C 26	2414.6	76.3	2814	6	AX454750	AX454750 Sequence
C 27	2414.6	76.3	2814	6	AX491228	AX491228 Sequence
C 28	2414.6	76.3	2814	6	AX658161	AX658161 Sequence
C 29	2414.6	76.3	2814	6	HSU32682	U73682 Human menin
C 30	2414.6	76.2	2865	9	BC038527	BC038527 Homo sapi
C 31	2410.6	76.2	169104	9	AL162377	AL162377 Human DNA
C 32	2403.2	75.2	2778	9	AF338233	AF338233 Homo sapi
C 33	2378	75.2	2778	9	BC043153	BC043153 Homo sapi
C 34	2340.2	74.0	2620	6	AX113851	AX113851 Sequence
C 35	2326.2	73.5	4345	9	AF273058	AF273058 Homo sapi
C 36	2326.2	73.5	4345	9	AL355585	AL355585 Human DNA
C 37	2302.2	72.8	76321	9	AC140149	AC140149 Homo sapi
C 38	2302.2	72.8	179950	2	AL162716	AL162716 Human DNA
C 39	2294.6	72.5	122916	9	HSM801357	BC051363 Homo sapi
C 40	2261.2	71.5	2664	9	BC051363	BC051363 Homo sapi
C 41	2247.4	71.0	2547	9	BC051363	AL158816 Homo sapi
C 42	2225.2	70.3	22885	2	AL158816	AF338234 Homo sapi
C 43	2207.2	69.8	4110	9	AF338234	AF338234 Homo sapi
C 44	2182.4	69.0	2651	9	HSM804728	AL833415 Homo sapi
C 45	2177.4	68.8	2601	9	BC031065	BC031065 Homo sapi

ALIGNMENTS

RESULT 1
AL135904/c

LOCUS

DEFINITION

Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains

part of the gene for a 7 transmembrane receptor (rhodopsin family)

(olfactory receptor like) protein, the gene for WGSCH-DJ0989615.3

protein, part of the PDPN3 (phosphodiesterase I/nucleotide

pyrophosphatase 3) gene, STSs and GSSs, complete sequence.

ACCESSION

AL135904

VERSION

AL135904.11

GI:7159399

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

59130 bp DNA linear PRI 20-JUL-2000
Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains
part of the gene for a 7 transmembrane receptor (rhodopsin family)
(olfactory receptor like) protein, the gene for WGSCH-DJ0989615.3
protein, part of the PDPN3 (phosphodiesterase I/nucleotide
pyrophosphatase 3) gene, STSs and GSSs, complete sequence.

Pred. No. is the number of results predicted by chance to have a

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

1 (bases 1 to 59130)
Lovell, J.
Direct Submission
Submitted (05-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7105768.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EMBL, ENBL; SWI, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
Chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP5-1005H11 is from the library RPCI-5 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: PCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP5-1005H11 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP5-1005H11 is at 59130 in this
sequence. The true right end of clone RP5-914N13 is at 105 in this
sequence.

FEATURES
Source

Location/Qualifiers

1..59130
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosomes="6"
/clone="RP5-1005H11"
/clone_lib="RPCI-5"

repeat_region
1..68
/note="AluS repeat: matches 1..68 of consensus"

repeat_region
546..844
/note="AluJo repeat: matches 2..301 of consensus"

repeat_region
870..940
/note="U4 repeat: matches 1..72 of consensus"

repeat_region
1691..2206
/note="L2 repeat: matches 1425..2737 of consensus"

repeat_region
3457..3960
/note="HAL1 repeat: matches 420..971 of consensus"

repeat_region
3971..4106
/note="MT1D repeat: matches 1..129 of consensus"

repeat_region
4449..4482
/note="17 copies 2 mer tg 100% conserved"

repeat_region
4966..5147
/note="AluJb repeat: matches 1..182 of consensus"

repeat_region
5216..5307
/note="L2 repeat: matches 2405..2503 of consensus"

repeat_region
5495..5794
/note="AluSx repeat: matches 1..296 of consensus"

repeat_region
5891..6180
/note="AluSx repeat: matches 1..291 of consensus"

repeat_region
7607..7618
/note="L1M1 repeat: matches 4856..4867 of consensus"

repeat_region
7619..8347
/note="L1P16 repeat: matches 5421..6156 of consensus"

repeat_region
8348..8898

/note="L1M1 repeat: matches 4866..5403 of consensus"
9018..9866
/note="TIGGER1 repeat: matches 1..836 of consensus"
9867..10189
/note="AluSx repeat: matches 1..309 of consensus"
10190..11723
/note="TIGGER1 repeat: matches 836..2418 of consensus"
12404..12715
/note="AluSx repeat: matches 1..312 of consensus"
complement(13947..14176)
/note="match: GSS: Em:B74942"
14280..14383
/note="AluSg/x repeat: matches 190..294 of consensus"
15340..15508
/note="L2 repeat: matches 2563..2750 of consensus"
15684..15968
/note="AluSx repeat: matches 35..312 of consensus"
16527..16612
/note="MIR repeat: matches 50..138 of consensus"
16731..17027
/note="AluSx repeat: matches 1..295 of consensus"
17609..17747
/note="MER91A repeat: matches 1..153 of consensus"
17867..18162
/note="AluY repeat: matches 1..299 of consensus"
18461..18950
/note="MER1A repeat: matches 2..527 of consensus"
19185..19224
/note="20 copies 2 mer tg 85% conserved"
19474..19712
/note="match: STS: Em:G16149"
20037..20361
/note="AluSx repeat: matches 1..311 of consensus"
complement(21566..22384)
/gene="dJ1005H11.1"
complement(21566..22384)
/gene="dJ1005H11.1"
/note="match: proteins: Sw:P47881 Tr:O70265 Tr:O70266
Tr:O70267 Tr:O70268 Tr:O62942 Tr:O95918 Tr:O95047"
/cdon start=1
/evidence="not experimental"
/product="dJ1005H11.1 (7 TRANSMEMBRANE RECEPTOR (RHODOP
FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN))"
/protein_id="CAB99212.1"
/db_xref="GI:9368991"
/db_xref="GOA:Q9NQN0"
/translation="LNGTILGLISLDRSLHAPMYFFLSHLAVDIAVACNTPR
NLLHPAKTSPAGRMQTFSTFAVTECLLVMSYDLVVAICHPFLAYLAIME
TILAVTSMTGVLISLIHLVLLPFCRPOKIYHFFCEILAVLKACADTHINE
LAGATSGLVPLSTIVSVYMCILCAIOIQSREYORKAFRTCSFLCVIGLYGT
MTVGRYGNPKSKYKLLIFSLFPLNPLICSRNSEVNTLKRVLGVERAL
complement(22229..22521)
/note="match: STS: Em:H39853"
22385..22489
/note="L2 repeat: matches 2461..2551 of consensus"
22696..23011
/note="AluSx repeat: matches 1..312 of consensus"
23044..23513
/note="match: STS: Em:AI066560"
complement(23123..23716)
/note="match: GSS: Em:AQ276140"
complement(23300..23543)
/note="match: GSS: Em:AQ058777"
23751..24305
/note="match: GSS: Em:AQ798220"
23753..24176
/note="match: GSS: Em:AQ550610"
23762..24485
/note="match: GSS: Em:B86674"
24220..24343
/note="FIAM C repeat: matches 1..124 of consensus"
complement(24681..25395)
/note="L1P16 repeat: matches 5421..6156 of consensus"

misc_feature /note="match: GSS: Em:AQ627093"
complement(24780..25383)
repeat_region /note="match: GSS: Em:AQ627184"
26053..26548
repeat_region /note="TIGER1 repeat: matches 1..521 of consensus"
26549..26861
repeat_region /note="LiM4 repeat: matches 4025..4339 of consensus"
26865..27121
repeat_region /note="Tiger3b repeat: matches 2..272 of consensus"
27122..27424
repeat_region /note="LiM4 repeat: matches 1..302 of consensus"
27425..27579
repeat_region /note="Tiger3b repeat: matches 272..451 of consensus"
27582..27658
repeat_region /note="Tiger3b repeat: matches 1155..1231 of consensus"
27659..29067
repeat_region /note="LiM4 repeat: matches 2585..4035 of consensus"
29189..29287
repeat_region /note="LiM4 repeat: matches 2499..2600 of consensus"
29316..29499
misc_feature /note="match: STS: Em:AA931824"
29507..29990
misc_feature /note="match: STS: Em:AA931824"
29530..29981
misc_feature /note="match: GSS: Em:AQ356756"
29530..29737
misc_feature /note="match: GSS: Em:AQ253729"
29533..29855
misc_feature /note="match: GSS: Em:AQ319835"
29538..29943
misc_feature /note="match: STS: Em:AA235111"
29560..29787
misc_feature /note="match: STS: Em:G27651 Em:T99793"
complement(29779..32112)
gene /gene="dJ1005H11.2"
CDS /gene="dJ1005H11.2"
/note="match: proteins: Tr:O95046"
/codon_start=1
/evidence=not experimental
/product="dJ1005H11.2 (WUGSC:H_DJ0988G:5.3 PROTEIN)"
Query Match 89.5%; Score 2832; DB 9; Length 59130;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;
QY 1 AGGTTTAAATCCATGAGACAGACGCAATTTTAAAGTGTTATTCACCAAAAATAAAGCTT 60
Db 32463 AGGTTTAAATCCATGAGACAGACGCAATTTTAAAGTGTTATTCACCAAAAATAAAGCTT 32404
QY 61 CAAATATGTGATGTGAAATCTGCCAGAACTAAGCGCGCGCGCTCAGACCCAGCGCTGC 120
Db 32403 CAAATATGTGATGTGAAATCTGCCAGAACTAAGCGCGCGCGCTCAGACCCAGCGCTGC 32344
QY 121 CTCAGGATGTAAGTGTAAACAAGGCGCCAGGGAGGTGTGGGGACACATGGGCGCTG 180
Db 32343 CTCAGGATGTAAGTGTAAACAAGGCGCCAGGGAGGTGTGGGGACACATGGGCGCTG 32284
QY 181 TGAGGCTGTGGGTGCCCGGTTCCCGAGTCCCGCCGAGCCCGCTCCACAGTGGTCCG 240
Db 32283 TGAGGCTGTGGGTGCCCGGTTCCCGAGTCCCGCCGAGCCCGCTCCACAGTGGTCCG 32224
QY 241 CTCGGTGTGTTCAGTGGCGGATTCGGGTTCAGACCCAGGCTCGGTTCGCCACC 300
Db 32223 CTCGGTGTGTTCAGTGGCGGATTCGGGTTCAGACCCAGGCTCGGTTCGCCACC 32164
QY 301 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCTCCGACGCTATGGAGGAG 360
Db 32163 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCTCCGACGCTATGGAGGAG 32104
QY 361 CTTGGTGTACCTCAGCCCTACTCGGGGTGTCTCTGGAGGAGCTACCGCAGTTGTG 420
Db 32103 CTTGGTGTACCTCAGCCCTACTCGGGGTGTCTCTGGAGGAGCTACCGCAGTTGTG 32044

QY 421 CGACCACTACCTGAGAGTATGAGACCAAGATCCTTATGGTTTCCATCGAACTG 480
Db 32043 CGACCACTACCTGAGAGTATGAGACCAAGATCCTTATGGTTTCCATCGAACTG 31984
QY 481 CTGTAATGTGAGCTGTTATTTGGATTTTGTGTTCTCCTTTTGTGAGAGTTT 540
Db 31983 GTGTAATGTGAGCTGTTATTTGGATTTTGTGTTCTCCTTTTGTGAGAGTTT 31924
QY 541 AGATCGGTTAGAGTCGGCTTTAGCTGGGAAGAGACAAATACTTGTGTCAACCTTTCT 600
Db 31923 AGATCGGTTAGAGTCGGCTTTATGTGGGAAGAGACAAATACTTGTGTCAACCTTTCT 31864
QY 601 GGACTAATTTGAAGAAAATGTAAACTACTTTGAAAAATTTAGCCTTATTTCAAAAAGATAT 660
Db 31863 GGACTAATTTGAAGAAAATGTAAACTACTTTGAAAAATTTAGCCTTATTTCAAAAAGATAT 31804
QY 661 GAAGCTATGAAGTAGAGTCATCTTTAGAGATGCCAGCTTTTGAGAGAGGCGGAGCAGAA 720
Db 31803 GAAGCTATGAAGTAGAGTCATCTTTAGAGATGCCAGCTTTTGAGAGAGGCGGAGCAGAA 31744
QY 721 GAAGCAGAACTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAATTTCTGAATCTTGAG 780
Db 31743 GAAGCAGAACTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAATTTCTGAATCTTGAG 31684
QY 781 GATGAATCCTCTCTGTAGAAAAAGAGCTTAAAAACAAGAGAAATCTAAACATTTCTCAACA 840
Db 31683 GATGAATCCTCTCTGTAGAAAAAGAGCTTAAAAACAAGAGAAATCTAAACATTTCTCAACA 31624
QY 841 GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAGTAGAGTCAAAATCC 900
Db 31623 GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAGTAGAGTCAAAATCC 31564
QY 901 CTCAAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATAGAGTGAAGAA 960
Db 31563 CTCAAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATAGAGTGAAGAA 31504
QY 961 CGACGGCTATAGCAATAAAAAAGATGCTTTGAATGAAAAATTTCAACTTCAGCAAGCCAT 1020
Db 31503 CGACGGCTATAGCAATAAAAAAGATGCTTTGAATGAAAAATTTCAACTTCAGCAAGCCAT 31444
QY 1021 AAAACAGCTTTTTCAGCAAGAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 1080
Db 31443 AAAACAGCTTTTTCAGCAAGAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 31384
QY 1081 CAGAAAAATAACATTTGAAGACTCCAAAGTACACGACAGAAACAAAGTTCTGAATGATAAGAA 1140
Db 31383 CAGAAAAATAACATTTGAAGACTCCAAAGTACACGACAGAAACAAAGTTCTGAATGATAAGAA 31324
QY 1141 AATCATCATCAGAACCTCGATGGAACAATTGCCAATGATGAAGATCAGGCTGCTGCTT 1200
Db 31323 AATCATCATCAGAACCTCGATGGAACAATTGCCAATGATGAAGATCAGGCTGCTGCTT 31264
QY 1201 GAAGAGACACAAACGATGATGAATAACCTGGAATTTAGAAGTGAACAGTGAATCGAAAAAT 1260
Db 31263 GAAGAGACACAAACGATGATGAATAACCTGGAATTTAGAAGTGAACAGTGAATCGAAAAAT 31204
QY 1261 GGTGTTTACTTAGATAATTCCTCCAAAAGAGGCTTTGAAGAACTGATTCATCTGCTGAAG 1320
Db 31203 GGTGTTTACTTAGATAATTCCTCCAAAAGAGGCTTTGAAGAACTGATTCATCTGCTGAAG 31144
QY 1321 TTAATAGCTTCTTTAAABACCTTAGAGGAGAAAGAACCAAAATTTATATTCAGTTGTCT 1380
Db 31143 TTAATAGCTTCTTTAAABACCTTAGAGGAGAAAGAACCAAAATTTATATTCAGTTGTCT 31084
QY 1381 GAAGTTGATAAAACAAAGAGAGCTTACAGAGCATATTTAAAAATTTTCAGACTCAACAA 1440
Db 31083 GAAGTTGATAAAACAAAGAGAGCTTACAGAGCATATTTAAAAATTTTCAGACTCAACAA 31024
QY 1441 GCATCTTTGCGATCAGAAAAACACATTTTGAATAATGAGATCAGAACTTCAACAGAAA 1500
Db 31023 GCATCTTTGCGATCAGAAAAACACATATTTTGAAGTGAAGATCAGAACTTCAACAGAAA 30964

repeat_region	7841..8087
	/rpt_family="L1"
repeat_region	8098..8672
	/rpt_family="L1"
repeat_region	10048..10070
	/rpt_family="AT_rich"
repeat_region	10071..10201
	/rpt_family="L1"
repeat_region	10472..10595
	/rpt_family="GA-rich"
repeat_region	10597..10624
	/rpt_family="(GGAA)n"
repeat_region	11168..11193
	/rpt_family="AT_rich"
repeat_region	11194..11732
	/rpt_family="L1"
repeat_region	11733..12143
	/rpt_family="MaLR"
repeat_region	12144..13177
	/rpt_family="L1"
repeat_region	13178..13549
	/rpt_family="MaLR"
repeat_region	13550..15665
	/rpt_family="L1"
repeat_region	15666..15929
	/rpt_family="Ricksha"
repeat_region	15928..17402
	/rpt_family="Ricksha"
repeat_region	17428..17608
	/rpt_family="L1"
repeat_region	17610..17974
	/rpt_family="L1"
repeat_region	18051..18316
	/rpt_family="L1"
repeat_region	18317..18619
	/rpt_family="Alu"
repeat_region	18620..19858
	/rpt_family="L1"
repeat_region	19869..19137
	/rpt_family="ERV1"
repeat_region	19241..19353
	/rpt_family="ERV1"
repeat_region	19354..19679
	/rpt_family="MER1_type"
repeat_region	19880..19999
	/rpt_family="ERV1"
repeat_region	21248..21680
	/rpt_family="L1"
repeat_region	21709..22051
	/rpt_family="MaLR"
repeat_region	23532..23933
	/rpt_family="MaLR"
repeat_region	25951..26252
	/rpt_family="Alu"
repeat_region	26503..26680
	/rpt_family="Alu"
repeat_region	27013..27185
	/rpt_family="MaLR"
repeat_region	27186..27549
	/rpt_family="MaLR"
repeat_region	27550..27672
	/rpt_family="MaLR"
repeat_region	27920..28076
	/rpt_family="ERV1"
repeat_region	31894..32190
	/rpt_family="MER2_type"
repeat_region	32191..32335
	/rpt_family="Alu"
repeat_region	32351..32441
	/rpt_family="MER2_type"
repeat_region	32609..32670
	/rpt_family="AT_rich"
repeat_region	32699..32819

```

repeat_region      /rpt_family="L1"
32820..33201
/rpt_family="MaLR"
33202..34039
/rpt_family="L1"
34104..34866
/rpt_family="L1"
37906..37943
/rpt_family="AT_rich"
37945..38043
/rpt_family="L1"
38044..38148
/product="snRNA-U6-related"
38044..38148
/rpt_family="U6"
38165..38548
/rpt_family="L1"
40180..40317
/rpt_family="L2"
40707..41010
/rpt_family="L2"
41410..41718
/rpt_family="Alu"

Query Match      89.5%; Score 2832; DB 9; Length 130030;
Best Local Similarity 94.5%; Pred No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;

QY 1 AGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 60
DB 34768 AGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 34827
QY 61 CAAATATGTGATGTGAAGTCCAGAACTTAAGCGCGCGGCTCAGACAGCGCTGC 120
DB 34828 CAAATATGTGATGTGAAGTCCAGAACTTAAGCGCGCGGCTCAGACAGCGCTGC 34887
QY 121 CTCAGGATGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGGCAACAATGGCCCTG 180
DB 34888 CTCAGGATGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGGCAACAATGGCCCTG 34947
QY 181 TGAGGCGCTGGGTGGCGCGGTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 240
DB 34948 TGAGGCGCTGGGTGGCGCGGTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 35007
QY 241 CTCGGGTGGTGTCTACGTGCGCATTCGGGTTCAGAGCCCAAGGCTGGTGTTCCTCCACC 300
DB 35008 CTCGGGTGGTGTCTACGTGCGCATTCGGGTTCAGAGCCCAAGGCTGGTGTTCCTCCACC 35067
QY 301 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGCAGCGCTCGACGCTATGGAGGAG 360
DB 35068 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGCAGCGCTCGACGCTATGGAGGAG 35127
QY 361 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGCGAGTGTGTG 35187
DB 35128 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGCGAGTGTGTG 35247
QY 421 GCAGCACTACTGAGAGTATGAGACAGATGAGAACTTATGGTTTTCATCGGAACTG 480
DB 35188 GCAGCACTACTGAGAGTATGAGACAGATGAGAACTTATGGTTTTCATCGGAACTG 35247
QY 481 GTGGTATGTCAGCTGTATTGATTTTGTGTTTCTCTTTTGTGGAGAGTGTG 540
DB 35248 GTGGTATGTCAGCTGTATTGATTTTGTGTTTCTCTTTTGTGGAGAGTGTG 35307
QY 541 AGATCGGTTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGGTGCACCGTTCT 600
DB 35308 AGATCGGTTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGGTGCACCGTTCT 35367
QY 601 GGACTAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCCTTATTCAAAAGAGTAT 660
DB 35368 GGACTAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCCTTATTCAAAAGAGTAT 35427
QY 661 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCCAGCTTTGAGAAGGAGGAGCAGAA 720

```

```

DB 35428 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCCAGCTTTGAGAAGCGCGCAGAA 35487
QY 721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGAGTCCAAATCTTGAATCTGAG 780
DB 35488 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGAGTCCAAATCTTGAATCTGAG 35547
QY 781 GATGAAATCCTCTGTCTAGAAAAGACCTTAAACAAGAGAAATCTTAAACATCTTCAACAA 840
DB 35548 GATGAAATCCTCTGTCTAGAAAAGACCTTAAACAAGAGAAATCTTAAACATCTTCAACAA 35607
QY 841 GATGAAATCCTCTGTCTAGAAAAGACCTTAAACAAGAGTATACAGTCTCTAGAGAAGTGAAGTCC 900
DB 35608 GATGAAATCCTCTGTCTAGAAAAGACCTTAAACAAGTATACAGTCTCTAGAGAAGTGAAGTCC 35667
QY 901 CTCAAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 960
DB 35668 CTCAAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 35727
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 1020
DB 35728 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 35787
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATCTTAAATAA 1080
DB 35788 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATCTTAAATAA 35847
QY 1081 CAGAAAATAACATTTTGAAGACTCCAAAGTACACGCAAGAACTCTGAAATGATAAAGAA 1140
DB 35848 CAGAAAATAACATTTTGAAGACTCCAAAGTACACGCAAGAACTCTGAAATGATAAAGAA 35907
QY 1141 AATCACATCAAGACCTGACTGGACACTTGCCTGCAATGATGAAAAGATCAGGCTGCTGCTT 1200
DB 35908 AATCACATCAAGACCTGACTGGACACTTGCCTGCAATGATGAAAAGATCAGGCTGCTGCTT 35967
QY 1201 GAAGAGACACACAGGATGATGATTAACCTGGAAATTAAGATGAACAGTGAATCGGAAAAAT 1260
DB 35968 GAAGAGACACACAGGATGATGATTAACCTGGAAATTAAGATGAACAGTGAATCGGAAAAAT 36027
QY 1261 GGTGCTTACTTAGATTAATCCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGCTGCTAAG 1320
DB 36028 GGTGCTTACTTAGATTAATCCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGCTGCTAAG 36087
QY 1321 TTAATGCTTTCTTTTAAACCTTTAGAGGAGAGAAACCAATTTTATATTCAGTTGTCT 1380
DB 36088 TTAATGCTTTCTTTTAAACCTTTAGAGGAGAGAAACCAATTTTATATTCAGTTGTCT 36147
QY 1381 GAAGTTTCAATAAACAAGAGAGAGCTTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 1440
DB 36148 GAAGTTTCAATAAACAAGAGAGAGCTTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 36207
QY 1441 GCATCTTTGAGTCAAGAAAACACACATTTTGAATAAGAGATCAGAGCTTCACACAGAAA 1500
DB 36208 GCATCTTTGAGTCAAGAAAACACATTTTGAATAAGAGATCAGAGCTTCACACAGAAA 36267
QY 1501 CTTAAAGTAAATGACTGAATTTATATCAAGAAAATGAAATGAAATCCACAGGAAATTAACA 1560
DB 36268 CTTAAAGTAAATGACTGAATTTATATCAAGAAAATGAAATGAAATCCACAGGAAATTAACA 36327
QY 1561 GTAGAGGAAATTTATCGGTTAGAGAAAAGAGAACTTTCTAAAGTGAATGAAGATC 1620
DB 36328 GTAGAGGAAATTTATCGGTTAGAGAAAAGAGAACTTTCTAAAGTGAATGAAGATC 36387
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
DB 36388 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGAAGAGAA 36447
QY 1681 TTGAGAGAACTATTTCATCTTATCAAGGGGAGATTTATTTCCNATGAGAAAAGAGCAT 1740
DB 36448 TTGAGAGAACTATTTCATCTTATCAAGGGGAGATTTATTTCCNATGAGAAAAGAGCAT 36507
QY 1741 GATTAATTCGTTGGAGCTCGGAATGCTGAAAAGAACTCAATGATTTTAAAGAAAGAAAT 1800

```

Db	36508	GATAATTGGTTGGCAGCTCGGACTGCTGAAAGAAACCTCAGTGATTTTAAAGAAAGAAAAT	36566
Qy	1801	GCTCAACAACAGACAAAATTAACATGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT	1860
Db	36568	GCTCAACAACAAAATTAACATGAAACAGAGTTCAAATTTGAACTTTTGAAGAAAAGAT	36627
Qy	1861	CCTTATGCACCTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATATGGTCCC	1920
Db	36628	CCTAATGCACCTCGATGTTTCAAATACAGCATTTGGCAGAGAGCATTTCCCATATATGGTCCC	36687
Qy	1921	TCACCAATTTGGTTGGCTTCATCTGAAACAAGAGCTTTTCTCTCTCTCTCCCAACTTTGTTG	1980
Db	36688	TCACCAATTTGGTTGGCTTTCACTGAAACAGAGCTTTTCCCTCTCTCTCCCAACTTTGTTG	36747
Qy	1981	GAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAAGGCTCACAGGC	2040
Db	36748	GAGGATCCACTCAGACTCTCACCTGTGCTTCCAGGGGAGGAGGAAGGCCCAAGCAGC	36807
Qy	2041	CCAGGGAATCTCTGGAACCATAGATTACAATGAAAGAGGAGAAATCAAGCTGTGATAGG	2100
Db	36808	CCAGGGAATCCCCGTGACCATCAGATTACCAATGAAAGAGGAGAACCAAGCTTATGACAGG	36867
Qy	2101	TTAAACCGATCCTCATAGGGTCCCTCTGACACTGGGTCCTCTGCTCACTCCATGGGACACAG	2160
Db	36868	TTAATCGATCCCTCACAGGGCTCCTCTGACACTGGGTCCCTGTCTCATCTCCGGTGGAAACAG	36927
Qy	2161	GACCGTAGATGATGTTTCTCTCGCAGAGACAATCATATCTCTGATTTAGCCCCCTCTCTCCA	2220
Db	36928	GACCGTAGATGATGTTTCTCTCACAGGGCAATCATATCTCTGATTTCAACTCTCTCTCCA	36987
Qy	2221	CAAAAGCAACACAGATTTTGTCTTAATTCCTGCTAGACTCTCTGGACCAAGCAGAACTCAGA	2280
Db	36988	CAAAAGGGAACACAGATTTTATTTCTAATTTCTGAAAGACTCTCTGGACCAGCAGAACCCAGA	37047
Qy	2281	AGTTTTAATATGCTTCTTTGGATAAAATGGATGGGTCAATGCGCTTCAGAAATGGAATCC	2340
Db	37048	AGTTTTAATAATGACTTCTTTGGATAAAATGGATAGGTCAATGCCCTTCAGAAATGGAATCC	37107
Qy	2341	AGTAGAAATGATACCAAAGATGATCTTTGTGTAAATTTAAATGTGCGCTGATTCATCTCCCT	2400
Db	37108	AGTAGAAATGATGCCAAAGATGATCTTTGTGTAAATTTAAATGTGCGCTGATTCATCTCCCT	37167
Qy	2401	GCTGAAATGAAGCCACTGGCCCTGGCTTTGTTCTCTCGACCTCTTGGCTCCAATCAGAGGT	2460
Db	37168	GCTGAAATGAAGCAACTGGCCCTGGCTTATTTCTCTCCACCTCTTGTCTCCAATCAGCGGT	37227
Qy	2461	CCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAGAGGACCTCTTTTCCCCCCA	2520
Db	37228	CCATTGTTTCCAGTGGATACAAGGGGCCCGTTTCATGAGAAGAGGACCTCTTTTCCCCCCA	37287
Qy	2521	CCCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAGGAGATTTCCCA	2580
Db	37288	CCCTCTCCAGGAACCATGTTTGGAGCTTCTCGAGGTTATTTTCCACCAGGAGATTTCCCA	37347
Qy	2581	GGTCCACCACTGCTCCATTGCAATGAGAAATGTCTATCCACCGAGGGGTTTCTCTCCT	2640
Db	37348	GGTCCACCACTGCTCCATTGCAATGAGAAAACATCTATCCACCGAGGGGTTTACCTCCT	37407
Qy	2641	TACCTTCCCCCAGACCTGATTTTTTCCGCCACCCGCACATCTTGAAGGTAGAATGAG	2700
Db	37408	TACCTTATCCGAGACCTGATTTTTTACCCCAACCCACATTTCTGAAGGTAGAAGCCAG	37465
Qy	2701	TTCCCTTCAGGTTTGATTCACCTTCAAATGAGCCTGCTACTGAAACATCCAGAAACACAG	2760
Db	37466	TTCCCTTCAGGATTTGATTTCCGCTTCAAAGGAGCCTGCTACTGGACATCCAGAAACACAG	37525
Qy	2761	CAGAGAACCTTGACATAATTTTCTCTCTCAAAAGTAAATTTTGATGATCTCATTTTCA	2820
Db	37526	CAGAGAACCTTGACATAATTTGCTTTCTTCAAAGTAAATTTTGATGATCTCATTTTCA	37585
Qy	2821	GTTTAAAGTAATGCTGTACTTAAGTGATTACATTTTCTCTCAAAATGGAAGCTTAATGGA	2880
Db	37586	GTTTAAAGTAATGCTGTACTTAAGTGATTGACATTTTCTCAAAATGGAAGTTAATGGA	37644

RESULT 3	AC005587	140915 bp	DNA	linear	PRI 07-NOV-2002
LOCUS	AC005587				
DEFINITION	AC005587	140915 bp	DNA	linear	PRI 07-NOV-2002
ACCESSION	AC005587				
VERSION	AC005587.1	GI:4156166			
KEYWORDS	HTG.				
SOURCE	HTG.				
ORGANISM	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.				
TITLE	HTG.				
JOURNAL	HTG.				
MEDLINE	HTG.				
REFERENCE	HTG.				
AUTHORS	HTG.			</	

Center code: WUGSC
 web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.edu
 ----- Summary Statistics -----
 Center project name: H_DJ0988G15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/csc>

SOURCE INFORMATION:
This clone was derived from human PAC library RPCL-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of AC005587;
actual end is at 140915 of AC005587.

FEATURES
SOURCE

RES	source	Location/Qualifiers
1.	.140915	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="16"
		/map="16"
		/clone="RPS-988G15"
		/clone_lib="RPC1-5"
repeat_region	13..2090	/rpt_family="L1"
repeat_region	2097..2299	/rpt_family="L1"
repeat_region	2300..2603	/rpt_family="Alu"
repeat_region	2604..4120	/rpt_family="L1"
repeat_region	4123..4166	/rpt_family="(TA)n"
repeat_region	4168..4307	/rpt_family="L1"
repeat_region	6635..6668	/rpt_family="AT-rich"
repeat_region	6829..7260	/rpt_family="L1"
repeat_region	7261..7642	/rpt_family="MALR"
repeat_region	7643..7923	/rpt_family="L1"
repeat_region	7982..8293	/rpt_family="Alu"
repeat_region	8286..8412	/rpt_family="L1"
repeat_region	8413..8708	

repeat_region	/rpt_family="Alu"	8709..8838	
repeat_region	/rpt_family="Li"	8840..8886	
repeat_region	/rpt_family="MaLR"	9111..9139	
repeat_region	/rpt_family="AT_rich"	9204..9490	
repeat_region	/rpt_family="Alu"	9646..9924	
repeat_region	/rpt_family="Alu"	10286..10455	
repeat_region	/rpt_family="L2"	10759..10838	
repeat_region	/rpt_family="Mariner"	11127..11366	
repeat_region	/rpt_family="L2"	11850..12154	
repeat_region	/rpt_family="Alu"	12172..12211	
repeat_region	/rpt_family="AT_rich"	12352..12469	
repeat_region	/rpt_family="(TAAG)n"	12458..12607	
repeat_region	/rpt_family="(CATA)n"	12585..12733	
repeat_region	/rpt_family="(TA)n"	12660..12777	
repeat_region	/rpt_family="(CATA)n"	12778..13031	
repeat_region	/rpt_family="Alu"	13226..13535	
repeat_region	/rpt_family="L2"	15279..15420	
repeat_region	/rpt_family="MIR"	15954..16022	
repeat_region	/rpt_family="MaLR"	16023..16332	
repeat_region	/rpt_family="MER2_type"	16333..16371	
repeat_region	/rpt_family="MaLR"	16630..16742	
repeat_region	/rpt_family="MaLR"	16816..17068	
repeat_region	/rpt_family="Li"	17088..17348	
repeat_region	/rpt_family="Li"	17857..17979	
repeat_region	/rpt_family="Alu"	18014..18506	
repeat_region	/rpt_family="MaLR"	18707..19081	
repeat_region	/rpt_family="MaLR"	19135..20717	
repeat_region	/rpt_family="Retroviral"	20728..20748	
repeat_region	/rpt_family="AT_rich"	21444..21599	
repeat_region	/rpt_family="MIR"	21652..22330	
misc_feature	/notes="similar to Mus musculus EST C88518 (NID:g3164280)"	21652..22254	
misc_feature	/notes="similar to Mus musculus EST W13172 (NID:g1287301) ma93h10.r1"	21652..22196	
misc_feature	/notes="similar to EST W23487 (NID:g1300321) zb33f06.r1"	21652..22127	
misc_feature	/notes="similar to Mus musculus EST W10529 (NID:g1284855) ma42b10.r1"	21657..22052	
misc_feature	/notes="similar to EST AA082271 (NID:g16244330) zn37g03.r1"	21657..22052	
misc_feature	/notes="similar to EST AA305238 (NID:g1957585)"	21657..22052	

```
misc_feature 21663..22255
/notes="similar to Mus musculus EST AA274895 (NID:g1914893)
va92c10.r1"
misc_feature 21665..22255
/notes="similar to EST AI097080 (NID:g3446662) oz22c05.x1"

Query Match      89.5%; Score 2832; DB 9; Length 140915;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;

QY 1 AGGTTTAAATCCATGAAGAAGACGACATTTTAAAGTGTAATCCACCAAAATAAAGCTT 60
Db 91632 AGGTTTAAATCCATGAAGAAGACGACATTTTAAAGTGTAATCCACCAAAATAAAGCTT 91691
QY 61 CAAATAATGTATGTGAAATCTCCAGAACTAAGCGCGCGGCTCAGACCAAGCGCTGC 120
Db 91692 CAAATAATGTATGTGAAATCTCCAGAACTAAGCGCGCGGCTCAGACCAAGCGCTGC 91751
QY 121 CTGAGGATGTAAAGTGTAAACAGAGGCCAGGGAGGTGTGGGGACACATGGGCCCTG 180
Db 91752 CTGAGGATGTAAAGTGTAAACAGAGGCCAGGGAGGTGTGGGGACACATGGGCCCTG 91811
QY 181 TGAAGCCTGTGGGTGCCCGGCTTCCCGAGCTCCCGCGCAGCCCGCTCCACAGTGGTCCG 240
Db 91812 TGAAGCCTGTGGGTGCCCGGCTTCCCGAGCTCCCGCGCAGCCCGCTCCACAGTGGTCCG 91871
QY 241 CTCGGTGTGTGTACGTGCGGATTCGGGTTCAGACCCCAAGCTGCGGTCTCCACC 300
Db 91872 CTCGGTGTGTGTACGTGCGGATTCGGGTTCAGACCCCAAGCTGCGGTCTCTCCACC 91931
QY 301 GCTTGTGTGGCCAGTGTACTCGGGTGACCGCCGACAGCAGCTCGACGCTATGGAGGAG 360
Db 91932 GCTTGTGTGGCCAGTGTACTCGGGTGACCGCCGACAGCAGCTCGACGCTATGGAGGAG 91991
QY 361 CTGGTGTCTACCTCAGCCCTACCTGGGCTGTCTCGAGGAGCTACCGAGATGTG 420
Db 91992 CTGGTGTCTACCTCAGCCCTACCTGGGCTGTCTCGAGGAGCTAGCCAGATGTG 92051
QY 421 GCACACTACTCTGAGAGTATGAGACCAAGATGAGAACTCTTATGTTTCCATCGGAACCTG 480
Db 92052 GCACACTACTCTGAGAGTATGAGACCAAGATGAGAACTCTTATGTTTCCATCGGAACCTG 92111
QY 481 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCCTTTTGTGGAGAGTTT 540
Db 92112 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCCTTTTGTGGAGAGTTT 92171
QY 541 AGATCGTTTAGAGTCCGCTTACCTGGAGAGAGCAAAACCTGGTGCAACGCTTCT 600
Db 92172 AGATCGTTTAGAGTCCGCTTATCTGGAGAGAGCAAAACCTGGTGCAACGCTTCT 92231
QY 601 GGACTAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCTTATTTCAAAAGAGTAT 660
Db 92232 GGACTAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCTTATTTCAAAAGAGTAT 92291
QY 661 GAAGCTATGAAGTAGAGTCACTTTAGAGATCCAGCTTGAAGAGGAGGAGGAGGAA 720
Db 92292 GAAGCTATGAAGTAGAGTCACTTTAGAGATCCAGCTTGAAGAGGAGGAGGAGGAA 92351
QY 721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 780
Db 92352 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 92411
QY 781 GATGAATCTCTGTCTAGAAAAGAGCTTAAACCAAGAGAACTTAACATTTCTCAACAA 840
Db 92412 GATGAATCTCTGTCTAGAAAAGAGCTTAAACCAAGAGAACTTAACATTTCTCAACAA 92471
QY 841 GATGAATGTAGGGGGATATTTCAAAAGATATACAGTCTCTAGAAGATGAGTCAAAATCC 900
Db 92472 GATGAATGTAGGGGGATATTTCAAAAGATATACAGTCTCTAGAAGATGAGTCAAAATCC 92531
QY 901 CTCAATCAAAAATAGCTGAAGCCMAATATCTGGAAGACATTTAAATAGAGTGAAGAA 960
Db 92532 CTCAATCAAAAATAGCTGAAGCCMAATATCTGGAAGACATTTAAATAGAGTGAAGAA 92591
```

```
QY 961 CGAGCGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCACAGCAAGCCAT 1020
Db 92592 CGAGCGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCACAGCAAGCCAT 92651
QY 1021 AAACAGCTTTTTCAGCAAGAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 1080
Db 92652 AAACAGCTTTTTCAGCAAGAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 92711
QY 1081 CAGAAAAATAACATTTGAAGACTCCAAAGTACACGCAGAACCAAGTTCTCAATGATAAGCAA 1140
Db 92712 CAGAAAAATAACATTTGAAGACTCCAAAGTACACGCAGAACCAAGTTCTCAATGATAAGCAA 92771
QY 1141 AATCATCTCAAGACCCCTGACTGGACACCTTGCCAAATGATGAAAGATCAGGCTGTGTGCTT 1200
Db 92772 AATCATCTCAAGACCCCTGACTGGACACCTTGCCAAATGATGAAAGATCAGGCTGTGTGCTT 92831
QY 1201 GAAGAAGACACAAACCGATGATGAATCTGGAATTTAGAAGTGAAACAGTGAATCGAAAAAT 1260
Db 92832 GAAGAAGACACAAACCGATGATGAATCTGGAATTTAGAAGTGAAACAGTGAATCGAAAAAT 92891
QY 1261 GGTGCTTACTTAGATPAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGTGAAG 1320
Db 92892 GGTGCTTACTTAGATPAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGTGAAG 92951
QY 1321 TTAAATGCTTCTTTAAACACCTTAGAGGAGAGAAACCAATTTATATTCAGTTGTCT 1380
Db 92952 TTAAATGCTTCTTTAAACACCTTAGAGGAGAGAAACCAATTTATATTCAGTTGTCT 93011
QY 1381 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAA 1440
Db 93012 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAA 93071
QY 1441 GCATCTTTGCGAGTCAAGAAACACACATTTTGAAGTGAAGTCAAGACTTCACAGAGAA 1500
Db 93072 GCATCTTTGCGAGTCAAGAAACACATATTTTGAAGTGAAGTCAAGACTTCACAGAGAA 93131
QY 1501 CTTAAAGTAAAGACTGAATTTATATCAAGAAATGAAATGAAATCTCAAGGAAATTAACA 1560
Db 93132 CTTAAAGTAAAGACTGAATTTATATCAAGAAATGAAATGAAATCTCAAGGAAATTAACA 93191
QY 1561 FTAGAGGAAATTTATCGGTTAGAGAAAGAGAAACTTTCTAAAGTGAAGTGAAGATC 1620
Db 93192 FTAGAGGAAATTTATCGGAAATAGAGAAAGAGAGACTTTCTAGAGTGAAGAAAGATC 93251
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGGCAAGATCTTTGAAGAGAA 1680
Db 93252 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCTAGCCAAAGATCTTTGAAGAGAA 93311
QY 1681 TTGGAGAGAACTATTCATCTTTATCAAGGCGAGATTTTCCATGAGAAAGAGCAAT 1740
Db 93312 TTGGAGAGAACTATTCATCTTTATCAAGGCGAGATTTTCCATGAGAAAGAGCAAT 93371
QY 1741 GATAATTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
Db 93372 GATAATTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 93431
QY 1801 GCTCAACACAGACAAATAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
Db 93432 GCTCAACACAGACAAATAATTAACCTGAAAGAGAGTTGAAATTTGAACCTTTTGAAGAAAGAT 93491
QY 1861 CTTTATGCACTGATGTTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATATCGTCCC 1920
Db 93492 CTTTATGCACTGATGTTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATATCGTCCC 93551
QY 1921 TCACCATTTGGTGGCCTTCATCTGAAAACAGAGCTTTTCTCTCTCTCCCAACTTTGTG 1980
Db 93552 TCACCATTTGGTGGCCTTCATCTGAAAACAGAGCTTTTCTCTCTCTCCCAACTTTGTG 93611
QY 1981 GAGGCTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
Db 93612 GAGGCTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 93671
```

QY	2041	CMAGGGAATCTCTGGAACATCAGATTACCAATGAAGAGGAGAAATCAAGCTGTGATAGG	2100
Db	93672	CCAGGGAATCCCTCGACCATCAGATTACCAATGAAGAGGAGAAACCAAGCTATGACAGG	93731
QY	2101	TTAAACCGATCTCATAGGGCTCCCTCTGACACTCGGTCTCTGTCACTCTCCATGCGAACGAC	2160
Db	93732	TTAAATCGATCTCACAGGGCTCCTTCTGACACTGGGTCCCTGTGATCTCCGGTGGAAACAG	93791
QY	2161	GACCGTAGATGATGTTTCTCTCGGACGGAACAATCATATCTGTATTGAGCCCTTCCTCCA	2220
Db	93792	GACCGTAGATGATGTTTCTCTCAACAGGGCAATCATATCTGTATTCAATCTTCCTCCA	93851
QY	2221	CAAAGGCAAGACAGATTTTGTGTTCTAATCTCGTAGACTGTCTGGACACGACGAATCTCAGA	2280
Db	93852	CAAAGGCAAGACAGATTTTATTCTAATCTGAAAGACTGTCTGGACACGACGAACCCAGA	93911
QY	2281	AGTTTAAATATGCTCTCTTTGGATATAATGATGGTCAATGCGCTTCAGAATCGAATCC	2340
Db	93912	AGTTTAAATATGACTTCTTTGGATATAATGATAGTCAATGCGCTTCAGAATCGAATCC	93971
QY	2341	AGTAGAAATGATACCAAGATGATCTTGTAATTTAAATGTGCGTGATTCATCTCTCCCT	2400
Db	93972	AGTAGAAATGATGCCAAAGATGATCTTGTAATTTAAATGTGCGTGATTCATCTCTCCCT	94031
QY	2401	GCTGAAATGAGCCACTGGCCCTGGCTTGTTCCTCCACCTCTGTGCTCCAAATCAGAGGT	2460
Db	94032	GCTGAAATGAGCAACTGGCCCTGGCTTATTTCTCCACCTCTTGCTCCAAATCAGCGGT	94091
QY	2461	CCATGTTTCCAGTGGATCCAGAGGCCCATCTTTGAGAAGGAGACCTCTTTCCGCCCA	2520
Db	94092	CCATGTTTCCAGTGGATCAAGGGGCCCATTCATGAGAAGGAGACCTCTTTCCGCCCA	94151
QY	2521	CCCTCTCCAGGAGCCATGTTTGGAGCTTCTCCAGATTAATTTTCCACCAAGGGAATTCCCA	2580
Db	94152	CCCTCTCCAGGAACATGTTTGGAGCTTCTCGAGGTATTTTCCACCAAGGGAATTCCCA	94211
QY	2581	GGTCCACCACTGCTCCATTTGCAATGAGAAATGTCATFCCACCGAGGGTITTTCTCTCT	2640
Db	94212	GGTCCACCACTGCTCCATTTGCAATGAGAAACATCTATCCACCGAGGGTITTTACCTCT	94271
QY	2641	TACCTTCCCCCAAGACTGGATTTTTCGCCCCACCCACCATCTCTGAAGGTAGAAAGTGAG	2700
Db	94272	TACCTTCATCCGAGACTGGATTTT--ACCCCAACCCACATCTCTGAAGGTAGAAAGGAG	94329
QY	2701	TTCCCTCAGGTTTGATTTCCACCTTCAAAATGAGCTGCTACTGAAACATCCAGAACCCACAG	2760
Db	94330	TTCCCTTCAGGATTTGATTTCCGCTTCAAAAGGAGCCTGCTACTGGACATCCAGAACCCACAG	94389
QY	2761	CAAGAAACCTGCAAAATATTTTGTCTCTCTCAAAAGTAATTTTGACTCATCTCAATTTC	2820
Db	94390	CAGACACCTGCAAAATATTTGTTGCTTCTTCAAAAGTAATTTTGACTCATCTCAATTTC	94449
QY	2821	GTTTAAAGTAACCTGCTTACTTAAGTGATTAACATTTTGTCTCAAAATGAAAGCTTAATGGA	2880
Db	94450	GTTTAAAGTAACCTGCTTACTTAAGTGATTAAGTACATTTT--CTCAAAATGAAAGTTAATGGA	94508
QY	2881	ATTATTAATCTCAGGATAGTATTTTGTGTAATAAGATGATTTTAAATATGAATCTTATGAG	2940
Db	94509	ATAATAGTTCTCAGGATAGTATTTTGTGTAATAAGATGGTTTGAATATGAATCTTATGAG	94568
QY	2941	TAAATTAATTTCAATTT-----TATTTTAGCGGTATACTATTTTCAAT--TTGATTAATC	2993
Db	94569	TAAATCAATTTTCCATTTTATATATCTAGATCATATACTTTAACTTGGTGAACATAATC	94628
QY	2994	CACATTAATATAACCAATAGTGGAGTTTATATATGTAATCTTTCAAGTGGGAGGCTT	3053
Db	94629	CACCTTAGAGAAACAATAGTGGAGTTTATATATGTAATCTTTCAGTGGGAGGCTT	94688
QY	3054	TAAATTTCTGAAGTCT---GGTCTTTTATGCCAAGAACTGTATTTACTGTGTGTGGACA	3110
Db	94689	TAAATTTCTGAAGGTTGTGGTGTCTTCATGCCAAGAACTGTATTTACTGTGTGTGTAGATA	94748
QY	3111	AATGTGAAGAACTGATTTATGCTTTAAATTAATTTATAGTTGATTTA 3154	

```

Db          94749 AATGTGAAGTAGACTTATGCTTAATTTAATAAATTTTCATGCA 94792
|||||
RESULT 4
AC004889/c
LOCUS      130030 bp      DNA      linear      PRI 02-OCT-2003
DEFINITION Homo sapiens PAC clone RP4-798C17 from 7, complete sequence.
ACCESSION  AC004889
VERSION     AC004889.1 GI:4156187
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 130030)
AUTHORS     Sulston,J.E. and Wilson,R.
TITLE        Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
PUBMED       9847074
REFERENCE   2 (bases 1 to 130030)
AUTHORS     Gibson,A. and Kozlowski.
TITLE        The sequence of Homo sapiens PAC clone RP4-798C17
JOURNAL      Unpublished (2001)
REFERENCE   3 (bases 1 to 130030)
AUTHORS     Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4 (bases 1 to 130030)
AUTHORS     Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5 (bases 1 to 130030)
AUTHORS     Waterston,R.
TITLE        Direct Submission
JOURNAL      Submitted (26-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6 (bases 1 to 130030)
AUTHORS     Wilson,R.
TITLE        Direct Submission
JOURNAL      Submitted (02-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jan 14, 1999 this sequence version replaced gi:3213108.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: sapiens@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_DJ0798C17
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NEGR1 Chromosome 7 Mapping

Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GrB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.

VECTOR: pCVPAC2.

FEATURES

source

Location/Qualifiers

1..130030
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"

/clone="RP4-798C17"

/clone_lib="RPCI-4"

5..3376

/rpt_family="L1"

3412..3820

/rpt_family="MER1_type"

4921..5004

/rpt_family="(TAGA)n"

5035..5081

/rpt_family="(TA)n"

7628..7696

/rpt_family="L1"

7697..7840

/rpt_family="(GAAA)n"

7841..8087

/rpt_family="L1"

8098..8672

/rpt_family="L1"

10048..10070

/rpt_family="AT-rich"

10071..10201

/rpt_family="L1"

10472..10595

/rpt_family="CA-rich"

10597..10624

/rpt_family="(GGAA)n"

11168..11193

/rpt_family="AT-rich"

11194..11732

/rpt_family="L1"

11733..12143

/rpt_family="MaLR"

12144..13177

/rpt_family="L1"

13178..13549

/rpt_family="MaLR"

13550..15665

/rpt_family="L1"

15666..15929

/rpt_family="Ricksha"

15928..17402

/rpt_family="Ricksha"

17428..17608

/rpt_family="L1"

17610..17974

/rpt_family="L1"

18051..18316

/rpt_family="L1"

18317..18619

/rpt_family="Alu"

repeat_region 18620..18858
/rpt_family="L1"
repeat_region 18869..19137
/rpt_family="ERVL"
repeat_region 19241..19353
/rpt_family="ERVL"
repeat_region 19354..19879
/rpt_family="MER1_type"
repeat_region 19880..19999
/rpt_family="ERVL"
repeat_region 21345..21680
/rpt_family="L1"
repeat_region 21709..22051
/rpt_family="MaLR"
repeat_region 23592..23933
/rpt_family="MaLR"
repeat_region 25951..26252
/rpt_family="Alu"
repeat_region 26503..26680
/rpt_family="Alu"
repeat_region 27013..27185
/rpt_family="MaLR"
repeat_region 27186..27549
/rpt_family="MaLR"
repeat_region 27550..27672
/rpt_family="MaLR"
repeat_region 27920..28076
/rpt_family="ERVL"
repeat_region 31894..32190
/rpt_family="MER2_type"
repeat_region 32191..32335
/rpt_family="Alu"
repeat_region 32361..32441
/rpt_family="MER2_type"
repeat_region 32605..32670
/rpt_family="AT-rich"
repeat_region 32699..32819
/rpt_family="L1"
repeat_region 32820..33201
/rpt_family="MaLR"
repeat_region 33202..34039
/rpt_family="L1"
repeat_region 34104..34866
/rpt_family="L1"
repeat_region 37906..37943
/rpt_family="AT-rich"
repeat_region 37945..38043
/rpt_family="L1"
snRNA 38044..38148
/product="snRNA-U6-related"
repeat_region 38044..38148
/rpt_family="U6"
repeat_region 38165..38548
/rpt_family="L1"
repeat_region 40186..40317
/rpt_family="L2"
repeat_region 40707..41010
/rpt_family="L2"
repeat_region 41410..41718
/rpt_family="Alu"

Query Match 89.5%; Score 2830.4; DB 9; Length 130030;

Best Local Similarity 94.5%; Pred. No. 0;

Matches 2990; Conservative 0; Mismatches 161; Indels 13; Gaps 5;

QY 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTTATTCACCAAAAAATAAGCTT 60

Db 121216 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTTATTCACCAAAAAATAAGCTT 121157

QY 61 CAAATATGTGATGTGAAACTGCAGAACTAAGCGCGCGGCTCAGACCGAGCTGC 120

Db 121156 CAAATATGTGATGTGAAACTGCAGAACTAAGCGCGCGGCTCAGACCGAGCTGC 121097

QY	121	CTCAGGATGTAAAGTGTAAACAAGAGGCCACAGGGAGGTGTGGGGGACAACATGGGCCTG	180
Db	121096	CTCAGGATGTGAAGTGTAAACAAGAGGCCACAGGGAGGTGTGGGGGACAACATGGGCCTG	121037
QY	181	TGAGGCCCTGTGGGTGCCCGGTTCGCCAGCTCCCGCAGCGCGTCCACAGTGTGTCG	240
Db	121036	TGAGGCCCTGTGGGTGCCCGGTTCGCCAGCTCCCGCAGCGCGTCCACAGTGTGTCG	120977
QY	241	CTCCGGTGTGTTGTACAGCTGCGCATTCGGGTTCAGACCCCAAGGCTGCGTGTTCTCCACC	300
Db	120976	CTCCGGTGTGTTGTACAGCTGCGCATTCGGGTTCAGACCCCAAGGCTGCGTGTTCTCCACC	120917
QY	301	GCTTGTGTGGCCAGTGTACTGGGTGACGCCAGACAGCCCTCGACGCTATGGAGGAG	360
Db	120916	GCTTGTGTGGCCAGTGTACTGGGTGACGCCAGACAGCCCTCGACGCTATGGAGGAG	120857
QY	361	CCTGGTGCTACCCCTCAGCCCTACCTGGGGTGTCTCTGGAGGAGCTACGACAGTTGTG	420
Db	120856	CCTGGTGCTACCCCTCAGCCCTACCTGGGGTGTCTCTGGAGGAGCTACGACAGTTGTG	120799
QY	421	GCAGCACTACTGAGAGTATGAGACCAAGATGAGAAATCCTTATGGTTTTCAATCGGAAC	480
Db	120796	GCAGCACTACTGAGAGTATGAGACCAAGATGAGAAATCCTTATGGTTTTCAATCGGAAC	120737
QY	481	GTGSTATGTGCAGCTGTATTGGATTTTTTGTGTCTCCTTTTGTGGAGAAGTTTT	540
Db	120736	GTGSTATGTGCAGCTGTATTGGATTTTTTGTGTCTCCTTTTGTGGAGAAGTTTT	120677
QY	541	AGATCGGTTAGGATCGGCTTACGTGGGAAGAGAGCAAAACCTTGTGCAACGCTTCT	600
Db	120676	AGATCGGTTAGGATCGGCTTACGTGGGAAGAGAGCAAAACCTTGTGCAACGCTTCT	120617
QY	601	GGACTAATTCGAAGAAAATGTAAACTACTCGAAAATTTAGCCCTTATTCAAAAAGAGTAT	660
Db	120616	GGACTAATTCGAAGAAAATGTAAACTACTCGAAAATTTAGCCCTTATTCAAAAAGAGTAT	120557
QY	661	GAAGGCTATGAAGTAGAGTCACTTTTAGAGGATGCCAGCTTTGAGAAGGAGCGACGAA	720
Db	120556	GAAGGCTATGAAGTAGAGTCACTTTTAGAGGATGCCAGCTTTGAGAAGGAGCGACGAA	120497
QY	721	GAAGCAGCAAGCTTTGGAGCAACCTGTCAAAGCTGACAGCTCCAAATCTCTGAACTTGAG	780
Db	120496	GAAGCAGCAAGCTTTGGAGCAACCTGTCAAAGCTGACAGCTCCAAATCTCTGAACTTGAG	120437
QY	781	GATGAAATCCTCTGTCTAGAAAAAGACTTAAAAAAGAGAAATCTAAACATTTCTCAACAA	840
Db	120436	GATGAAATCCTCTGTCTAGAAAAAGACTTAAAAAAGAGAAATCTAAACATTTCTCAACAA	120377
QY	841	GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAGATGAGTCAAAATCC	900
Db	120376	GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAGATGAGTCAAAATCC	120317
QY	901	CTCAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA	960
Db	120316	CTCAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA	120257
QY	961	CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATCTCAACTTCAGACAAAGCCAT	1020
Db	120256	CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATCTCAACTTCAGACAAAGCCAT	120197
QY	1021	AAACAGCTTTTTTCAGCAGAAGCTGAAGTATCGAAAAGGAGAGTGCAGTGAACTTAATAAA	1080
Db	120196	AAACAGCTTTTTTCAGCAGAAGCTGAAGTATCGAAAAGGAGAGTGCAGTGAACTTAATAAA	120137
QY	1081	CAGAAAATAACATTTGAAGACTCCAAAGTACGCGAGAACAGAGTTCTGTAATGATAAGAA	1140
Db	120136	CAGAAAATAACATTTGAAGACTCCAAAGTACGCGAGAACAGAGTTCTGTAATGATAAGAA	120077
QY	1141	AATCACATCAAGACCTGTGACACTTGCACATGATCAAAAGATCAGAGCTGCTGTGCTT	1200
Db	120076	AATCACATCAAGACCTGTGACACTTGCACATGATCAAAAGATCAGAGCTGCTGTGCTT	120017
QY	1201	GAAGAGACACACCGATGATATACCTGGAATTTAGAAGTGAACAGTGAATGGGAAAT	1260

120016	DB	GAAGAGACACAAACGGATGATGATACCTTGGAAATTAGAAGTGAACAAGTCAATGGGAAT	119957
1261	QY	GGTGCCTTACTTAGATAAATCCTCCAAAGGAGCTTTGAAGAAACTGATTCATGCTGCTAAG	1320
119956	DB	GGTGCTAACTTAGATGATCCTCTGAAAGAGCTTTTGAAGAAACTGATTCATGCTGCTAAG	119897
1321	QY	TTAAATGCTCTTTAAAAACCTTAGAAGAGAGAAAGAAACCAATTTATATTCAGTTGTCT	1380
119896	DB	TTAAATGTTCTTTAAAAAGCTTAGAAGGAGAAAGAAACCAATTTATTCAGTTATCT	119837
1381	QY	GAAATTGATAAAACAAAGGAAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	1440
119836	DB	GAAATGGACAAAACAAAGGAAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	119777
1441	QY	GAATCTTTGCAGTCAGAAAACACACATTTTGAATAATGAGAAATCAGAGCTTCAACAGAAA	1500
119776	DB	GCACTCTTTGCAATTCAGAAAACATATTTTGAAGTGAAGAAATCAGAGCTTCAACAGAAA	119717
1501	QY	CTTAAAGTAAATGACTGAAATATATCAAGAAAAATGAAATGAAATCTTACAGAAATTAACA	1560
119716	DB	CTTAAATAATGACTGAAATCTATCAAGAAATGAAATGAAATCTTACAGAAATTAACA	119657
1561	QY	GTAGAGAAATATTCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGATC	1620
119656	DB	GTGAGGAAATATCCGAATAGAGGAAAGAGAGAGCTTTCTAGAGTGGAGAAAGATC	119597
1621	QY	AGCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCAAGAGATCTTGAAGAGAA	1680
119596	DB	AGCGTGCCACTGAAGGCTGGAGACCTATAGAAAGCTAGCCAAAGATCTTGAAGAGAA	119537
1681	QY	TTGAGAGAACTATTCAATCTTATCAAGGCGAGATTAATTTCCCATGAGAAAAAGACAT	1740
119536	DB	TTGAGAGAACTGTTCATTTTATCAAAAGCAGGTTATTTCTTACGAGAAAGAGGACAT	119477
1741	QY	GATTAATGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT	1800
119476	DB	GATAATTTGGTTGGCAGCTCGGACTGCTGAAAGAAACCTCAGTGATTTAAGGAAAGAAAT	119417
1801	QY	GCTCACAAACGACAAAAATTAACATGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAGAT	1860
119416	DB	GCTCACAAACAAACAAAAATTAACATGAAACAGAGTTGAAATTTGAACCTTTTGAAGAAAGAT	119357
1861	QY	CCTTATGCACTCGATGTTCCAAATACAGCATTTGCGAGAGAGCATTCGCCATATGCTGCC	1920
119356	DB	CCTAATGCACTCGATGTTCCAAATACAGCATTTGCGAGAGAGCATTCGCCATATGCTGCC	119297
1921	QY	TCACCATTTGGTTGGCTTTCATCTGAAACAGAGCTTTCTCTCTCTCCCAACTTTGTTG	1980
119296	DB	TCACCATTTGGTTGGCTTTCATCTGAAACAGAGCTTTCTCTCTCTCCCAACTTTGTTG	119237
1981	QY	GAGGTCACCTCAGACTCTACCTTTCCTCCAGGGGAGAGGAGGCTCAGAGGC	2040
119236	DB	GAGGATCCACTCAGACTCTACCTTTCCTCCAGGGGAGAGGAGGCTCAGAGGC	119177
2041	QY	CCAGGATCTCTGGACCATCAGATTACCAATGAAAGAGAGAAATCAAGCTGTGATAGG	2100
119176	DB	CCAGGAAATCCCTGGACCATCAGATTACCAATGAAAGAGAGAAATCAAGCTGTGATAGG	119117
2101	QY	TTAAACCGATCTCATAGGGCTCCCTCTGACATGGGTCTCTGTACCTCCATGGGACGAG	2160
119116	DB	TTAAATCGATCTCAGAGGCTCCTTCTGACATGGGTCCCTGTCTATCCGCTGGAAACAG	119057
2161	QY	GACCGTAGAGTAGTGTTCCTCCGCGACGACAAATCATATCCTGATTCAGCCCTCTCTCCA	2220
119056	DB	GACCGTAGAGTAGTGTTCCTCCACACAGGCAATCATATCCTGATTCACCTCTCTCTCCA	118997
2221	QY	CAAAGCGAAGACAGATTTTGTTCATTTCTGGTAGACTGTCTGGACACAGAGAACTCAGA	2280
118996	DB	CAAAGGGAAGACAGATTTTATCTAAATCTGAAGAGACTGTCTGGACACAGAGAACTCAGA	118937
2281	QY	AGTTTAAATAGCTTCTTTGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAATCC	2340

Db 118936 AGTTTAAATGACTTCTTTGGTAAATGAGTAGGTCAATGCCCTTCAGAAATGGATCC 118877
Qy 2341 AGTAGAAATGATACCAAGATGATCTGTGTAATTTAAATGTGCGCTGATTCATCTCTCCCT 2400
Db 118876 AGTAGAAATGATGCCAAGATGATCTGTGTAATTTAAATGTGCGCTGATTCATCTCTCCCT 118817
Qy 2401 GCTGAAATGAAGCCACTGSCCTGCTTGTTCCTCCACCTCTTCTCTCCAAATCAGAGGT 2460
Db 118816 GCTGAAATGAAGCACTGSCCTGCTTGTTCCTCCACCTCTTCTCTCCAAATCAGAGGT 118757
Qy 2461 CCATTGTTCCAGTGATGCAAGAGGCCCAATCTTGAGAGGAGCACTCTCTTCCGCCCA 2520
Db 118756 CCATTGTTCCAGTGATGCAAGAGGCCCAATCTTGAGAGGAGCACTCTCTTCCGCCCA 118697
Qy 2521 CTTCTCCAGAGGCAATGTTTGGAGTCTTCGAGATATTTTCCACCAAGGATTTCCCA 2580
Db 118696 CTTCTCCAGAGGCAATGTTTGGAGTCTTCGAGATATTTTCCACCAAGGATTTCCCA 118637
Qy 2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGTTCCTCTCT 2640
Db 118636 GGTCCACCACTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGTTCCTCTCT 118577
Qy 2641 TACTTCCCAAGACTGATTTTCCGCCACCCCACTTCTCAAGGTGAAGTGA 2700
Db 118576 TACTTCCCAAGACTGATTTTCCGCCACCCCACTTCTCAAGGTGAAGTGA 118519
Qy 2701 TTTCCCTCAGGTTGATTCACCTTCAATGAGCCTGCTACTGAAATCAGAACCCACAG 2760
Db 118518 TTTCCCTCAGGTTGATTCGCTTCAAGAGGCTGCTACTGAAATCAGAACCCACAG 118459
Qy 2761 CAAGAAACCTGACAAATTTTCTCTCTTCAAAAGTAATTTGACTGATCTCATTTTCA 2820
Db 118458 CAAGAAACCTGACAAATTTTCTCTCTTCAAAAGTAATTTGACTGATCTCATTTTCA 118399
Qy 2821 GTTTAAGTAATGCTGTACTTAAGTGATTAACCTTTTGTCTCAATTTGAAGCTTAATGGA 2880
Db 118398 GTTTAAGTAATGCTGTACTTAAGTGATTAACCTTTTGTCTCAATTTGAAGCTTAATGGA 118340
Qy 2881 ATTATAATCTCAGGATGATTTTGAATAAAGATGATTTAAATGATGATCTTATGAG 2940
Db 118339 ATATAGTCTCAGGATGATTTTGAATAAAGATGATTTGAATGATGATCTTATGAG 118280
Qy 2941 TAAATTTTCAATTTTATTTAGAGGATTAACCTTTTGAATTTCAATTTTGAATTAATC 2993
Db 118279 TAAATCAATTTTCAATTTTATTTATTTAGATCATATTAACCTTTTGAATTTCAATTT 118220
Qy 2994 CACTATTATATAAATAGTGGAGTTTATATATATATATATATATATATATATATATATAT 3053
Db 118219 CACTCTTAGAGAAACATAGTGGAGTTTATATATATATATATATATATATATATATATAT 118160
Qy 3054 TAAATTTCTGAGTCTTGTCTTTATGCAAGAACTGATTTTACTGTGTTGTGACA 3110
Db 118159 TAAATTTCTGAGTCTTGTCTTTATGCAAGAACTGATTTTACTGTGTTGTGACA 118100
Qy 3111 AATGTGAAGTAACTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3154
Db 118099 AATGTGAAGTAACTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 118056

RESULT 5
AC144412/c
LOCUS
DEFINITION Homo sapiens chromosome UNK clone RP11-224H1, WORKING DRAFT
SEQUENCE, 2 unordered pieces.
AC144412
VERSION AC144412.1 GI:29824566
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157801)
AUTHORS Waterston, R.H.

The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 157801)
Waterston, R.H.
Direct Submission
Submitted (12-APR-2003) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H.NH0224H01
----- Summary Statistics -----
Sequencing vector: M13; %
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 158459 bases at least Q40
Consensus quality: 158679 bases at least Q30
Consensus quality: 158958 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 27114: contig of 27114 bp in length
* 27115 157801: gap of unknown length
* 27215 157801: contig of 130587 bp in length.

FEATURES
source
1..157801
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clone="RP11-224H1"
1..27114
/note="assembly_name:Contig13"
27215..157801
/note="assembly_name:Contig14"

misc_feature
misc_feature
ORIGIN

Query Match 88.6%; Score 2804.2; DB 2; Length 157801;
Best Local Similarity 94.2%; Pred No. 0;
Matches 2981; Conservative 0; Mismatches 163; Indels 19; Gaps 6;
Qy 1 AGGTTTAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCCACCAAAATTAAGCTT 60
Db 79615 AGGTTTAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCCACCAAAATTAAGCTT 79556
Qy 61 CAAATATGTGATGTGAATACTGCCAGAACTAAGCGCGCGGCTCAGACCAGCGCTGC 120
Db 79555 CAAATATGTGATGTGAATACTGCCAGAACTAAGCGCGCGGCTCAGACCAGCGCTGC 79496
Qy 121 CTCAGATGATAAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGACAACAATGGGCTG 180
Db 79495 CTCAGATGATAAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGACAACAATGGGCTG 79436
Qy 181 TGAGGCTGTGGTGCCTCCGCTTCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCTCG 240
Db 79435 TGAGGCTGTGGTGCCTCCGCTTCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCTCG 79376
Qy 241 CTCGCGTGTGGTGTCACTGCGCATTCGGGTTCAGACCCCAAGGCTGTGTCTCCACC 300

79375 CTCCGGTGGTGTCTACGTCGCCGATTCGGGTTCACAGCCCAAGGCTGGTGTCTCCACC 79316
QY
301 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCAGAGAGAGCTCGACGCTATGAGGAG 360
Db
79315 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCAGAGAGAGCTCGACGCTATGAGGAG 79256
QY
361 CTTGTGTCTACCCCTCAGCCCTACCTGCGGTGCTGCTGAGAGAGCTAGCAGAGTGTG 420
Db
79255 CCGGTGTCTACCCCTCAGCCCTACCTGCGGTGCTGCTGAGAGAGCTAGCAGAGTGTG 79196
QY
421 GCAGCACTACCTGAGAGTATGAGACAGATGAGATCCTTATGTTTTCATCGGAACTG 480
Db
79195 GCAGCACTACCTGAGAGTATGAGACAGATGAGATCCTTATGTTTTCATCGGAACTG 79136
QY
481 GTGGTATGTCAGCTGTATGAGATTTTGTGTTTCTCTCTTTTGTGAGAGAGTGT 540
Db
79135 GTGGTATGTCAGCTGTATGAGATTTTGTGTTTCTCTCTTTTGTGAGAGAGTGT 79076
QY
541 AGATCGGTAGGAGTGGCTTTACGTGGGAGAGAGAGCAAACTTGGTGCACGCTTCT 600
Db
79075 AGATCGGTAGGAGTGGCTTTACGTGGGAGAGAGAGCAAACTTGGTGCACGCTTCT 79016
QY
601 GGAATAATGAGAGAAATGTAACACTTGAATAATTTAGCTTATTCAAAAGAGTAT 660
Db
79015 GGAATAATGAGAGAAATGTAACACTTGAATAATTTAGCTTATTCAAAAGAGTAT 78956
QY
661 GAAGGCTATGAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 720
Db
78955 GAAGGCTATGAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 78896
QY
721 GAAGCAGAGTGTGGAGCAACCTCTGAAAGCTGAAAGCTGCAATTTCTGAACTTGG 780
Db
78895 GAAGCAGAGTGTGGAGCAACCTCTGAAAGCTGAAAGCTGCAATTTCTGAACTTGG 78836
QY
781 GATGAATCTCTGTCTAGAGAAAGCTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 840
Db
78835 GATGAATCTCTGTCTAGAGAAAGCTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 78776
QY
841 GATGAATCTCTGTCTAGAGAAAGCTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 900
Db
78775 GATGAATCTCTGTCTAGAGAAAGCTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 78716
QY
901 CTCAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATAGTAGAGAGAA 960
Db
78715 CTCAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATAGTAGAGAGAA 78656
QY
961 CGACGGCTATAGCAATTAAGATGCTTGAATGAAATTTCACTTCCAGACAGCCAT 1020
Db
78655 CGACGGCTATAGCAATTAAGATGCTTGAATGAAATTTCACTTCCAGACAGCCAT 78596
QY
1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTGAATTAATAA 1080
Db
78595 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTGAATTAATAA 78536
QY
1081 CAGAAAATACATTTGAGAGCTCCAAAGTACACGAGAGAGAGTCTGATGATAGAGAA 1140
Db
78535 CAGAAAATACATTTGAGAGCTCCAAAGTACACGAGAGAGAGTCTGATGATAGAGAA 78476
QY
1141 AATCAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTGAATTAATAA 1200
Db
78475 AATCAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTGAATTAATAA 78416
QY
1201 GAAGAGACACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db
78415 GAAGAGACACACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 78356
QY
1261 GGTGCTTTTCTAGATATCTCTCAAAAGAGAGCTTTGAAGAGAACTGATGATGATGATG 1320
Db
78355 GGTGCTTTTCTAGATATCTCTCAAAAGAGAGCTTTGAAGAGAACTGATGATGATGATG 78296
QY
1321 TTAATGCTTTCTTAAAGAGCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db
78295 TTAATGCTTTCTTAAAGAGCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 78236

1381 GAAGTTGATAAAAAAGAGAGCTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 1440
Db
78235 GAAGTTGATAAAAAAGAGAGCTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 78176
QY
1441 GCATCTTTGAGTCAAGAAACACACATTTTGAATAATGAGATCAGAGCTTCAACAGAA 1500
Db
78175 GCATCTTTGAGTCAAGAAACACATATTATTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 78116
QY
1501 CTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db
78115 CTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 78056
QY
1561 GTAGAGAAATTTATCGTTTGAAGAAAGAGAGAACTTTTAAAGTATGATGATGATGATGATG 1620
Db
78055 GTAGAGAAATTTATCGTTTGAAGAAAGAGAGAACTTTTAAAGTATGATGATGATGATGATG 77996
QY
1621 AGCCATCCCTCAGAGAGCTGAGAGCTATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db
77995 AGCCATCCCTCAGAGAGCTGAGAGCTATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 77936
QY
1681 TTGAGAGAGAGTATTTTCAATTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db
77935 TTGAGAGAGAGTATTTTCAATTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77876
QY
1741 GATTAATTTGTTGTCAGCTGCGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db
77875 GATTAATTTGTTGTCAGCTGCGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77816
QY
1801 GCTCACAACAG 1860
Db
77815 GCTCACAACAG 77756
QY
1861 CCATTTGTCAGCTGAG 1920
Db
77755 CCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 77696
QY
1921 TCACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1980
Db
77695 GCACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 77636
QY
1981 GAGGTCAGCTCAGAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2040
Db
77635 GAGGTCAGCTCAGAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 77576
QY
2041 CCAGGAGTCTCTGAG 2100
Db
77575 CCAGGAGTCTCTGAG 77516
QY
2101 TTAACCGATCTCTAGAGGCTCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db
77515 TTAACCGATCTCTAGAGGCTCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77456
QY
2161 GACCGTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db
77455 GACCGTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 77396
QY
2221 CAAAGGAG 2280
Db
77395 CAAAGGAG 77336
QY
2281 AGTTTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db
77335 AGTTTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 77276
QY
2341 AGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Db
77275 AGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 77216
QY
2401 GCTGAAGATGAG 2460
Db
77215 GCTGAAGATGAG 77156

QY 2461 CCATGTTTCCAGTGTGATGACAGAGGCCCATCTTTCGAGAGGAGCCTCTTTCCTCCCA 2520
 Db 7155 CCATGTTTCCAGTGTGATGACAGAGGCCCATCTTTCGAGAGGAGCCTCTTTCCTCCCA 77096
 QY 2521 CTTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGAGATTCCCA 2580
 Db 77095 CTTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGAGATTCCCA 77036
 QY 2581 GGTCCACACCTGCTCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCT 2640
 Db 77035 GGTCCACACCTGCTCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCT 76976
 QY 2641 TACTTTCCTCCCAAGACTGGAATTTTCCCTCCCAAGGAGCCTCTTTCCTCCCA 2700
 Db 76975 TACTTTCCTCCCAAGACTGGAATTTTCCCTCCCAAGGAGCCTCTTTCCTCCCA 76918
 QY 2701 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCCTGCTACTGAAATCAATCAGAACACAG 2760
 Db 76917 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCCTGCTACTGAAATCAATCAGAACACAG 76866
 QY 2761 CAGAAACCTGACAAATTTTCTCTCTTCAAAAGTAATTTGACTGATCTCATTTTCA 2820
 Db 76865 CAGAAACCTGACAAATTTTCTCTCTTCAAAAGTAATTTGACTGATCTCATTTTCA 76806
 QY 2821 GTTAAAGTAACTGCTGTTTAAAGTAACTGTTTAAAGTAACTGTTTAAAGTAACTG 2880
 Db 76805 GTTAAAGTAACTGCTGTTTAAAGTAACTGTTTAAAGTAACTGTTTAAAGTAACTG 76746
 QY 2881 ATTAAATCTCAGGATGATGATTTTCTTAAATGAGTAAATTAATGATGATGATG 2940
 Db 76745 ATTAAATCTCAGGATGATGATTTTCTTAAATGAGTAAATTAATGATGATGATG 76686
 QY 2941 TAAATTAATTTCAATTTTAAATTAAGGATGATGATTTTAAATTAAGGATGATG 2994
 Db 76685 TAAATTAATTTCAATTTTAAATTAAGGATGATGATTTTAAATTAAGGATGATG 76626
 QY 2995 ACTAATTAATTAAGGATGATGATTTTAAATTAAGGATGATGATTTTAAATTAAGG 3054
 Db 76625 ACTAATTAATTAAGGATGATGATTTTAAATTAAGGATGATGATTTTAAATTAAGG 76566
 QY 3055 AAATTTCTGAAGTCTTCTCTCTTAAATTAAGGATGATGATTTTAAATTAAGGATG 3111
 Db 76565 AAATTTCTGAAGTCTTCTCTCTTAAATTAAGGATGATGATTTTAAATTAAGGATG 76506
 QY 3112 ATGTGAAGTAACTTATGCTTAAATTAAGTAACTTATGATTTA 3154
 Db 76505 ATGTGAAGTAACTTATGCTTAAATTAAGTAACTTATGATTTA 76463

RESULT 6
 AC106862/c
 LOCUS AC106862 169723 bp DNA linear PRI 15-OCT-2003
 DEFINITION Homo sapiens BAC clone RP11-61L23 from 7, complete sequence.
 ACCESSION AC106862
 VERSION AC106862.3 GI:19310327
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 169723)
 AUTHORS Sulston, J.E. and Wilson, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 95063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 169723)
 AUTHORS Bradshaw-Cordum, H., Elliott, G., Meyer, R. and Spalding, L.
 TITLE The sequence of Homo sapiens BAC clone RP11-61L23
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 169723)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Submitted (12-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 169723)
 Waterston, R.H.
 Direct Submission
 Submitted (09-MAR-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 169723)
 Waterston, R.H.
 Direct Submission
 Submitted (13-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 6 (bases 1 to 169723)
 Wilson, R.
 Direct Submission
 Submitted (15-OCT-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 9, 2002 this sequence version replaced gi:18250112.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0061L23

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa K., Woon, P.Y., Zhao, B., Frengen, E.,
 Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

FEATURES
 source
 1..169723
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="RP11-61L23"
 /clone_lib="RPC1-11"
 389..525
 /rpt_family="ERV1"
 939..1317
 repeat_region
 repeat_region

repeat_region	18250..18317	/rpt_family="TC2"	
repeat_region	18318..18624	/rpt_family="Alu"	
repeat_region	18625..18682	/rpt_family="TC2"	
repeat_region	19381..20275	/rpt_family="Alu"	
repeat_region	20765..21158	/rpt_family="ERV1"	
repeat_region	24151..24238	/rpt_family="MIR"	
repeat_region	25086..25464	/rpt_family="L2"	
repeat_region	27302..27400	/rpt_family="MIR"	
repeat_region	28301..28509	/rpt_family="MIR"	
repeat_region	28733..23749	/rpt_family="L1"	
repeat_region	30312..30547	/rpt_family="MaLR"	
repeat_region	30697..30990	/rpt_family="Alu"	
repeat_region	31833..31907	/rpt_family="MER1_type"	
repeat_region	32075..32197	/rpt_family="L2"	
repeat_region	32305..32561	/rpt_family="Alu"	
repeat_region	34217..34583	/rpt_family="MaLR"	

Query Match 88.6%; Score 2804.2; DB 9; Length 169723;

Best Local Similarity 94.2%; Pred. No. 0;

Matches 2961; Conservative 0; Mismatches 163; Indels 19; Gaps 6;

QY	1	AGGTTTAATCATGAAGAAGCAGCAATTTTAAAGTGTATTACCAAAAATTAAGCTT	60
Db	57205	AGGTTTAATCATGAAGAAGCAGCAATTTTAAAGTGTATTACCAAAAATTAAGCTT <td>57146</td>	57146
QY	61	CAAAATATGTGATGTGAATTCGCAGAACTGCGGCGCGCGGCTCAGACCGCGCTGC	120
Db	57145	CAAAATATGTGATGTGAATTCGCAGAACTGCGGCGCGCGGCTCAGACCGCGCTGC	57086
QY	121	CTCAGATGTAAAGTGTAAAGAGCGCGCGGAGGTGTGGGGACACATGGGCTG	180
Db	57085	CTCAGATGTAAAGTGTAAAGAGCGCGGAGGTGTGGGGACACATGGGCTG	57026
QY	181	TGAGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCGCGCGCTCCACAGTGGTCCG	240
Db	57025	TGAGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCGCGCGCTCCACAGTGGTCCG	56966
QY	241	CTCCGGTGTGGTGTACGTGGCGATTCCGGTTCAGACCCGAAGCTCGTGTCTCCACC	300
Db	56965	CTCCGGTGTGGTGTACGTGGCGATTCCGGTTCAGACCCGAAGCTCGTGTCTCCACC	56906
QY	301	GCTTGTGTGGCGAGTGTACTGCGGTGACCGCCAGACAGCAGCTCGAGGAG	360
Db	56905	GCTTGTGTGGCGAGTGTACTGCGGTGACCGCCAGACAGCAGCTCGAGGAG	56846
QY	361	CTGTGTCTACCCCTCAGCCCTACCTGGGCTGTCTGGAGGAGCTACGAGAGTCTG	420
Db	56845	CTGTGTCTACCCCTCAGCCCTACCTGGGCTGTCTGGAGGAGCTACGAGAGTCTG	56786
QY	421	GCAGCACTACCTGAGAGTATGAGACCATGAGATCCCTTATGGTTTTCCATCGGAATG	480
Db	56785	GCAGCACTACCTGAGAGTATGAGACCATGAGATCCCTTATGGTTTTCCATCGGAATG	56726
QY	481	GTGGTATGTGACGTGTATTCGATTTTTTGTGTTCTCCCTTTTTTGTGAGAGATTTT	540
Db	56725	GTGGTATGTGACGTGTATTCGATTTTTTGTGTTCTCCCTTTTTTGTGAGAGATTTT	56666

QY 541 AGATCGGTTAGGAGTCGGCTTTACGTTGGAGAGACGCAAAACTTGGTGCACGCTTCT 600
Db 5665 AGATCGGTTAGGAGTCGGCTTTACGTTGGAGAGACGCAAAACTTGGTGCACGCTTCT 5666
QY 601 GGAATAATTGAAGAAAATGTAACACTACTTGAATAATTTAGCCTTATTTCAAAAGAGTAT 660
Db 56605 GGAATAATTGAAGAAAATGTAACACTACTTGAATAATTTAGCCTTATTTCAAAAGAGTAT 56546
QY 661 GRAGGCTATGAGTAGAGTCATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCGACAGAA 720
Db 56545 GAAGGCTATGAGTAGAGTCATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCGACAGAA 56486
QY 721 GAAGCAGCAAGTTTGGAGGCAACCTCTGAAAAGCTGAAACAGGTCCTCAATTTCTGAACTTGAG 780
Db 56485 GAAGCAGCAAGTTTGGAGGCAACCTCTGAAAAGCTGAAACAGGTCCTCAATTTCTGAACTTGAG 56426
QY 781 GATGAATCTCTGCTAGAAAAGACTTTAAACAGAGGAAATCTAAACATTTCTCAAGAA 840
Db 56425 GATGAATCTCTGCTAGAAAAGACTTTAAACAGAGGAAATCTAAACATTTCTCAAGAA 56366
QY 841 GATGAATCTGATGCGGATATTTCAAAAGATATACAGTCTCTAGAAGATGAGTCAAAATCC 900
Db 56365 GATGAATCTGATGCGGATATTTCAAAAGATATACAGTCTCTAGAAGATGAGTCAAAATCC 56306
QY 901 CTCGAATCACAATAGCTGAAGCAAAATCACTGCAAGACATTTAAATGATGAGAA 960
Db 56305 CTCGAATCACAATAGCTGAAGCAAAATCACTGCAAGACATTTAAATGATGAGAA 56246
QY 961 CGACGGCTATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAGGCCAT 1020
Db 56245 CGACGGCTATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAGGCCAT 56186
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
Db 56185 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 56126
QY 1081 CAGAAAATAACATTTGAAGACTCCAAAGTACACGACAGAACAGTCTCGAATGATAAGAA 1140
Db 56125 CAGAAAATAACATTTGAAGACTCCAAAGTACACGACAGAACAGTCTCGAATGATAAGAA 56066
QY 1141 AATCAATCAAGACCTTGACTGACACTTGCCAAATGATGAAGATCAGGCTGCTGCTT 1200
Db 56065 AATCAATCAAGACCTTGACTGACACTTGCCAAATGATGAAGATCAGGCTGCTGCTT 56006
QY 1201 GRAGAGACACACGAGATGATTAACCTGGATTTAGAGTGAACAGTGAATCGGAAAT 1260
Db 56005 GRAGAGACACACGAGATGATTAACCTGGATTTAGAGTGAACAGTGAATCGGAAAT 55946
QY 1261 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 55945 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 55886
QY 1321 TTAATGCTTTCTTAATAACCTTAGAGGAGAGAAACCAATTTATATTCAGTTGCTCT 1380
Db 55885 TTAATGCTTTCTTTAAAGGCTTAGAGGAGAGAAACCAATTTATATTCAGTTGCTCT 55826
QY 1381 GAAGTTGATAAAACAAAGGAAGAGCTTACAGAGCATATTTAAATCTTCAGACTCAACAA 1440
Db 55825 GAAGTTGATAAAACAAAGGAAGAGCTTACAGAGCATATTTAAATCTTCAGACTCAACAA 55766
QY 1441 GCATCTTGCAGTCAGAAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAAA 1500
Db 55765 GAATCTTTGCAATCAGAAAACATATATTTTGAAGTGAAGATCAGAGCTTCAACAGAAA 55706
QY 1501 CTTAAAGTAAATGACTGAATATATCAAGAAAATGAAATGAACTCCACAGGAAATTAACA 1560
Db 55705 CTTAAABATATGACTGAATCTATCAAGAGATGAATGAACTCTATAGGAATTAACA 55646
QY 1561 GTAGAGGAAAATTTTCGGTTAGAGAAAAGAGAGAACTTTCTAAAGTAGATGAAAAGATC 1620
Db 55645 GTAGAGGAAAATTTCCGAATAGAGAAAAGAGAGAGCTTTCTAAAGTGAAGAAAAGCTC 55586
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGCAAGAGATCTTGAAGAGAA 1680

Db 55585 AGCCGTCCTCACTGAAACAGCTGGAGACCTATAGAAAAGCTACCAAGATCTTGAAGAGAA 55526
QY 1681 TTGAGAGAACTATCTATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAGAGCAT 1740
Db 55525 TTGAGAGAACTGTTCTATTTTATCAAAAGCAGGTATTTCTTACGAGAAAAGAGCAT 55466
QY 1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTTAAAGGAAAAGAAAAT 1800
Db 55465 GATAATTTGGTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTTAAAGGAAAAGAAAAT 55406
QY 1801 GCTCACAACAGACAAAATTAACCTGAAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT 1860
Db 55405 GCTCACAACAGACAAAATTAACCTGAAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT 55346
QY 1861 CTTATGCACTCGATGTTTCCAAATACAGATTTTGGCAGAGAGATTTCCCATATGCTGCTCC 1920
Db 55345 CTTATGCACTGATGTTTCCAAATACAGATTTTGGCAGAGAGATTTCCCATATGCTGCTCC 55286
QY 1921 TCACATTTGGTGGCTTCACTGAAAACAGAGCTTTTCTCTCTCTCCAACTTTGTTG 1980
Db 55285 GCACCATTTGGTGGCTTCACTGAAAACAGAGCTTTTCTCTCTCTCCAACTTTGTTG 55226
QY 1981 GAGGTCCTCACTCAGACTCTCACCTTTCTCCAGGGGAGAGAGAGGCTCACGAGGC 2040
Db 55225 GAGATCCACTGGGACTCTCACCTGCTTCCGAGGAGAGAGAGGCCCCAAGAGGC 55166
QY 2041 CCAGGGAATCCTTGGACCAATCAGATTAACCAATGAAAAGAGAGAAATCAAGCTGTAAGG 2100
Db 55165 CCAGGGAATCCTTGGACCAATCAGATTAACCAATGAAAAGAGAGAAATCAAGCTGTAAGG 55106
QY 2101 TTAACCGATCCTCATAGGCTTCCCTGACACTGGGCTCTGTCCACCTCCATGAGGACAG 2160
Db 55105 TTAACCGATCCTCATAGGCTTCCCTGACACTGGGCTCTGTCCACCTCCATGAGGACAG 55046
QY 2161 GACCGTAGAGTATGTTTCTCCGCGAGGACCAATCATATCCTGATTCAGCCCTTCTCTCCA 2220
Db 55045 GACTGAAGATGATGTTTCTCCGCGAGGACCAATCATATCCTGATTCAGCTCTCTCTCCT 54986
QY 2221 CAAAGGCAAGACATTTGTTCTTAATCTGGTAGACTGTGTGACAGCAGAGACTCAGA 2280
Db 54985 CAAAGGCAAGACATTTGTTCTTAATCTGGTAGACTGTGTGACAGCAGAGACTCAGA 54926
QY 2281 AGTTTAAATATGCTTCTTTGGATAAAATGATGGGTCAATGCCCTTCAGAAAATGGAATCC 2340
Db 54925 AGTTTAAATATGCTTCTTTGGATAAAATGATGGGTCAATGCCCTTCAGAAAATGGAATCC 54866
QY 2341 AGTAGAAATGATACCAAGAGATGATCTGGTAATTTAAATGTGCTGATTCATCTCTCCT 2400
Db 54865 AGTAGAAATGATACCAAGAGATGATCTGGTAATTTAAATGTGCTGATTCATCTCTCCT 54806
QY 2401 GCTGAAAATGAAGCACTGCGCTTGTCTTCTCCACCTTGTCTTCTCAATCAGAGGT 2460
Db 54805 GCTGAAAATGAAGCACTGCGCTTGTCTTCTCCACCTTGTCTTCTCAATCAGAGGT 54746
QY 2461 CAATTTCTTCAAGTGAAGAGGCCCCATCTTGAAGAGAGGAGGCTCTTTTCCCCCA 2520
Db 54745 CAATTTCTTCAAGTGAAGAGGCCCCATCTTGAAGAGAGGAGGCTCTTTTCCCCCA 54686
QY 2521 CCTCTCTCAGAGGAGGCTTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGAGATTTCCA 2580
Db 54685 CCTCTCTCAGAGGAGGCTTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGAGATTTCCA 54626
QY 2581 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCATTCACCGAGGGGTTTCTCTCT 2640
Db 54625 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCATTCACCGAGGGGTTTCTCTCT 54566
QY 2641 TACCTTCCCAAGACCTGGATTTTCCCCCCCCACCATTTCTGAAGGTAGAGTGAAG 2700
Db 54565 TACCTTCCCAAGACCTGGATTTTCCCCCCCCACCATTTCTGAAGGTAGAGTGAAG 54508
QY 2701 TTTCCCTCAGGTTTGAATTCACCTTTCAATGAGGCTGCTACTGAAACATCTCAGAACCA 2760


```
Db 54507 TTCCTTCAGGATTGATTCGCCCTTAA-----TGCTACTGAACATCCAGGACCACA- 54456
QY 2761 CAGAAACCTGACATATTTTCTCTCTCTCAAAAGTAATTTTCTGACTGATCTCATTTTCA 2820
Db 54455 AAAGAAACCTGACATATTTGCTCTCTCTCAAAAGTAATTTTCTGACTGATCTCATTTTCA 54396
QY 2821 GTTAAAGTAACCTGCTGTTTACTTAAGTGAATACATTTTGTCTCAAAATGAAGCTTAATGGA 2880
Db 54395 GTTAAAGTAACCTGCTGTTTACTTAAGTGAATACATTTTGTCTCAAAATGAAGCTTAATGGA 54336
QY 2881 ATTATAATCTCAGCATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 2940
Db 54335 ATTATAATCTCAGCATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 54276
QY 2941 TAAATTAATTTCAATTTTATTTTACAGGTATATCTAATTTCAATTTG-----ATTAAATCC 2994
Db 54275 TAAATCAATTTCAATTTTATTTTCTAGATCATATACTATTTAATTTTGGTGAATTAATCC 54216
QY 2995 ACTATATATAAACAATAGTGGAGTATTTATATATGTAATCTTTTCCAGTGGGGGCTTT 3054
Db 54215 ACTGTATATAAACAATAATAGGAGTATTTATATATGTAATCTTTTCCAGTGGGGGCTTT 54156
QY 3055 AAATCTGAAGTCT---GTCGCTTTATGTCGAAGAACTGTATTTACTGTGTTGTGACAA 3111
Db 54155 AAATCTGAAGTCTGTTGGTGTCTTCATGCAAGAACTGTATTTACTGTGTTGTAGATAA 54096
QY 3112 ATGTGAAGTAATCTTATGCTTAATAATAATTTAGTTGATTTA 3154
Db 54095 ATGTGAAGTAATCTTATGCTTAATAATAATTTAGTTGATTTA 54053

RESULT 7
AC143343 175228 bp DNA linear HTG 06-APR-2003
LOCUS Homo sapiens chromosome UNK clone RP11-513B18, WORKING DRAFT
DEFINITION SEQUENCE, 3 unordered pieces.
ACCESSION AC143343
VERSION AC143343.1 GI:29569308
KEYWORDS HTG; HTGS; PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175228)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
2 (bases 1 to 175228)
AUTHORS Waterston,R.H.
REFERENCE Direct Submission
TITLE Submitted (06-APR-2003) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H NH0513B18
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 17422 bases at least Q40
Consensus quality: 174667 bases at least Q30
Consensus quality: 174887 bases at least Q20
* NOTE: This is a 'working draft' sequence. It currently
```

```
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1545: contig of 1545 bp in length
* 1546 1545: gap of unknown length
* 1646 46992: contig of 45347 bp in length
* 46993 47092: gap of unknown length
* 47093 175228: contig of 128136 bp in length.
*
* Location/Qualifiers
* 1 175228
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="UNK"
* /clone="RP11-513B18"
* /note="assembly_name:Contig8"
* /note="assembly_name:Contig9"
* /note="assembly_name:Contig10"
*
misc_feature
1..1545
/notes="assembly_name:Contig8"
misc_feature
1646..46992
/notes="assembly_name:Contig9"
misc_feature
47093..175228
/notes="assembly_name:Contig10"

ORIGIN
Query Match 88.6%; Score 2804.2; DB 2; Length 175228;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 2981; Conservative 0; Mismatches 163; Indels 19; Gaps 6;

QY 1 AGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 60
Db 77979 AGTTTAAATTCATGAAGAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 78038
QY 61 CAAATATGATGATGTAATACTCCAGACTAGAGCGGCGCGGCTCAGACCAAGCGCTGC 120
Db 78039 CAAATATGATGATGTAATACTCCAGACTAGAGCGGCGGCGGCTCAGACCAAGCGCTGC 78098
QY 121 CTCAGGATGTAAGTGTAAACAAGAGCGCCAGGAGGAGTGTGGGGGACAACTGGGCTG 180
Db 78099 CTCAGGATGTAAGTGTAAACAAGAGCGCCAGGAGGAGTGTGGGGGACAACTGGGCTG 78158
QY 181 TGAGGCGCTGTGGTGGCGCGGCTTCCAGCTCCCCCGAGCGCGCTCCACAGTGTGCGG 240
Db 78159 TGAGGCGCTGTGGTGGCGCGGCTTCCAGCTCCCCCGAGCGCGCTCCACAGTGTGCGG 78218
QY 241 CTCGCGTGTGTTGTACGTGCGCAATTCGGTTCCAGACCAAGCGTCCGTTCTCCACC 300
Db 78219 CTCGCGTGTGTTGTACGTGCGCAATTCGGTTCCAGACCAAGCGTCCGTTCTCCACC 78278
QY 301 GCTTGTGTGGCCAGTGTATTCTCGGTGACCGCCAGAGAGCGCTCCAGCGTATGGAGGAG 360
Db 78279 GCTTGTGTGGCCAGTGTATTCTCGGTGACCGCCAGAGAGCGCTCCAGCGTATGGAGGAG 78338
QY 361 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCTGGAGAGCTACCGAGTTGTG 420
Db 78339 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCTGGAGAGCTACCGAGTTGTG 78398
QY 421 GCAGCACTACTGAGAGATGAGACCAAGATGAGATCTTTATGTTTCCATCGGACTG 480
Db 78399 GCAGCACTACTGAGAGATGAGACCAAGATGAGATCTTTATGTTTCCATCGGACTG 78458
QY 481 GTGGTATGTGAGCTGTATTGGATTTTTTGTGTTCTCTCTTTTGTGGAGAGTTT 540
Db 78459 GTGGTATGTGAGCTGTATTGGATTTTTTGTGTTCTCTCTTTTGTGGAGAGTTT 78518
QY 541 AGATCGGTAGGATGCGCTTTACGTGGGAAGAGAGCAAAAACCTGGTGCACGCTTCT 600
Db 78519 AGATCGGTAGGATGCGCTTTACGTGGGAAGAGAGCAAAAACCTGGTGCACGCTTCT 78578
QY 601 GGACTAATTAAGAAATGTAATACTACTTGAAATTTAGCTTATTATTAAGAGATAT 660
```



```

QY 2821 GTTTAAGTAACCTGCTGTTACTTAACTGATTAACACCTTTTGCCTCAAATGAAGCTTAATGGA 2880
Db 80789 GTTTAAGTAACCTGCTGTTACTTAACTGATTAACACCTTTTGCCTCAAATGAAGCTTAATGGA 80848
QY 2881 ATTATAAATCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 2940
Db 80849 ATTATAAATCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 80908
QY 2941 TAAATATTTCAAAT--TTATTTTAGACGGTATATACTATTTCAATTTG---ATTAATCC 2994
Db 80909 TAAATCAATCCATTTTATTTATTTCTAGATCAATATACTATTTTAAATTTGGTGAATTAATCC 80968
QY 2995 ACTATTATATAACATATAGTGGAGCTTTATATATATGTAATCTTTCAGGTGGGAGGCTTT 3054
Db 80969 ACTGTTATAGAACAAATATAGGAGCTTTTATATATGTAATCTTTCAGGTGGGAGGCTTT 81028
QY 3055 AAATCTGTAAGTCT---GTGCTCTTTATGCCAAGAACTGATTTACTGTGTTGTGGACAA 3111
Db 81029 AAATCTGAAGGTTGTGGTGTCTTCATGCCAAGAACTGATTTACTGTGTTGTAGATAA 81088
QY 3112 ATGTGAAGTAACCTTATGCTTTAAATATAATATATAGTATTA 3154
Db 81089 ATGTGAAGTAACCTTATGCTTTAAATATAATATAGTATTA 81131

RESULT 8
AC073264 200123 bp DNA linear PRI 15-OCT-2003
LOCUS Homo sapiens BAC clone RP11-298A10 from 7, complete sequence.
DEFINITION AC073264
ACCESSION AC073264.5 GI:13443255
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200123)
AUTHORS Sulston,J.E. and Wilson,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 200123)
AUTHORS Abbott,A., Doebber,A., Martinka,S. and Tomlins,C.
TITLE The sequence of Homo sapiens BAC clone RP11-298A10
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 200123)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 200123)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 200123)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 200123)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 24, 2001 this sequence version replaced gi:11192200.
----- Genome Center
COMMENT Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu

```

Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0298A10

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GR7/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Taneno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is AC024155. Actual start of this clone is at base position 1 of RP11-298A10 actual end is at base position 200123 of RP11-298A10.

FEATURES	source
repeat_region	1. 200123 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="7" /map="7"
repeat_region	/clone="RP11-298A10" /clone_lib="RPCI-11" 259..763 /rpt_family="L1"
repeat_region	779..1347 /rpt_family="ERV1"
repeat_region	1345..2101 /rpt_family="ERV1"
repeat_region	2321..2728 /rpt_family="ERV1"
repeat_region	2745..2945 /rpt_family="L1"
repeat_region	3043..3082 /rpt_family="A-rich"
repeat_region	3102..3293 /rpt_family="L2"
repeat_region	3340..3452 /product="rRNA-28S_rRNA-related"
repeat_region	3478..4173 /rpt_family="L2"
repeat_region	5403..5702 /rpt_family="Alu"
repeat_region	6528..6566 /rpt_family="L2"

```

repeat_region 7084. .7363
/rpt_family="Alu"
repeat_region 8128. .8452
/rpt_family="MIR"
repeat_region 8613. .9033
/rpt_family="MaLR"
repeat_region 9081. .9193
/rpt_family="L1"
repeat_region 10068. .10096
/rpt_family="(CA)n"
repeat_region 10121. .10721
/rpt_family="ERV1"
repeat_region 10788. .10943
/rpt_family="MIR"
repeat_region 11143. .11902
/rpt_family="ERV1"
repeat_region 11348. .12146
/rpt_family="MaLR"
repeat_region 12147. .12832
/rpt_family="L1"
repeat_region 12833. .13301
/rpt_family="MaLR"
repeat_region 13390. .13824
/rpt_family="MaLR"
repeat_region 14968. .15129
/rpt_family="MIR"
repeat_region 15148. .15245
/rpt_family="L1"
repeat_region 15277. .15433
/rpt_family="L1"
repeat_region 15434. .15745
/rpt_family="Alu"
repeat_region 15746. .16513
/rpt_family="L1"
repeat_region 16521. .16612
/rpt_family="MIR"
repeat_region 17075. .17710
/rpt_family="L2"
repeat_region 17711. .18004
/rpt_family="Alu"
repeat_region 18005. .18067
/rpt_family="L2"
repeat_region 18068. .18390
/rpt_family="Alu"
repeat_region 18391. .18442
/rpt_family="L2"
repeat_region 18929. .19301
/rpt_family="MaLR"
repeat_region 19952. .20213
/rpt_family="L2"
repeat_region 20215. .20243
/rpt_family="(TGAA)n"
repeat_region 20312. .20346
/rpt_family="AT-rich"
repeat_region 20566. .21492
/rpt_family="L1"
repeat_region 21493. .22063
/rpt_family="ERVX"
repeat_region 22064. .24908
/rpt_family="L1"
repeat_region 24909. .25219
/rpt_family="Alu"
repeat_region 25220. .26128
/rpt_family="L1"
repeat_region 26226. .26550
/rpt_family="MaLR"
repeat_region 28925. .29206
/rpt_family="Alu"
repeat_region 29821. .30421
/rpt_family="L1"
repeat_region 30504. .30842
/rpt_family="L1"
repeat_region 30846. .30930

```

```

/rpt_family="MaLR"
repeat_region 30938. .31014
/rpt_family="MIR"
repeat_region 32007. .32232
/rpt_family="L1"
repeat_region 32232. .32696
/rpt_family="L1"
repeat_region 32697. .32995
Query Match 88.6%; Score 2804.2; DB 9; Length 200123;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 2981; Conservative 0; Mismatches 163; Indels 19; Gaps 6;
QY 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 60
Db 120786 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 120845
QY 61 CAAATAATGTGATGTGAAAACTGCCAGAACTAAAGCGGGCGGGCTCAGACCGGCTGC 120
Db 120846 CAAATAATGTGATGTGAAAACTGCCAGAACTAAAGCGGGCGGGCTCAGACCGGCTGC 120905
QY 121 CTCAGGATGTTAAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGACAAACATGGGCTG 180
Db 120906 CTCAGGATGTTAAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGACAAACATGGGCTG 120965
QY 181 TGAGCCCTGTGGGTGCCCGCGCTTCCCAGCTCCCGCCGAGCCCGCTCCACAGTGGTCCG 240
Db 120966 TGAGCCCTGTGGGTGCCCGCGCTTCCCAGCTCCCGCCGAGCCCGCTCCACAGTGGTCCG 121025
QY 241 CTCGGTTGGTGTTCACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTGTCTCCACC 300
Db 121026 CTCGGTTGGTGTTCACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTGTCTCCACC 121085
QY 301 GCTTGTTCGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCCTCGACGCTATGGAGGAG 360
Db 121086 GCTTGTTCGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCCTCGACGCTATGGAGGAG 121145
QY 361 COTGCTGTACCCCTCAGCCCTACTCGGGCTGTCTGGAGGAGCTACGACAGAGTTGTG 420
Db 121146 CCGGTGTACCCCTCAGCCCTACTCGGGCTGTCTGGAGGAGCTACGACAGAGTTGTG 121205
QY 421 GCAGCACTACCTGAGAGTATGAGACCAATCCTTAAGTGTTCATCGGAATCG 480
Db 121206 GCAGCACTACCTGAGAGTATGAGACCAATCCTTAAGTGTTCATCGGAATCG 121265
QY 481 GTGGTATCGCAGCTGTTATTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 121266 GTGGTATCGCAGCTGTTATTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121325
QY 541 AGATCGGTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGTGCAACGCTTCT 600
Db 121326 AGATCGGTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGTGCAACGCTTCT 121385
QY 601 GGACTAATGAAGAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT 660
Db 121386 GGACTAATGAAGAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT 121445
QY 661 GAAGGCTATGAAGTAGAGTCATCTTTAGAGTAGCCAGCTTTGAGAGGAGGAGCAGAA 720
Db 121446 GAAGGCTATGAAGTAGAGTCATCTTTAGAGTAGCCAGCTTTGAGAGGAGGAGCAGAA 121505
QY 721 GAAGCAGAGAGTTTGGAGGCAACCTGTGAAAGCTGAACAGTCCCAATCTGAACCTGAG 780
Db 121506 GAAGCAGAGAGTTTGGAGGCAACCTGTGAAAGCTGAACAGTCCCAATCTGAACCTGAG 121565
QY 781 GATGAATCCTCTGTCTAGAAAAGACCTTAAACAAGAGAAATCTAAACATTTCTCAACA 840
Db 121566 GATGAATCCTCTGTCTAGAAAAGACCTTAAACAAGAGAAATCTAAACATTTCTCAACA 121625
QY 841 GATGAATGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 900
Db 121626 GATGAATGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 121685

```

QY	901	CTCAAAATCAAAATAGCTGAAGCCAAATATCATCTGCAAGACATTTTAAAAATGAGTGAAGAA	960
Db	121686	CTCAAAATCACAAATAGCTGAAGCCAAATATCATCTGCAAGACATTTTAAAAATGAGTGAAGAA	121745
QY	961	CGACGGCTATAGCAATAAAAGATGCTTTTGAATGAAAAATTTCTCAAATTCAGACAAAGCCAT	1020
Db	121746	CGACGGCTATAGCAATAAAAGATGCTTTTGAATGAAAAATTTCTCAAATTCAGACAAAGCCAT	121805
QY	1021	AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAACCTTAATATAA	1080
Db	121806	AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAACCTTAATATAA	121865
QY	1081	CAGAAAAATAACATTTTGAAGACTCCAAAGTACACGCAGAACAAAGTTCTGAATGATATAAGAA	1140
Db	121866	CAGAAAAATAACATTTTGAAGACTCCAAAGTACACGCAGAACAAAGTTCTGAATGATATAAGAA	121925
QY	1141	AATCACATCAAGACCGCTGACTGGACACTTGCCTATGATGAAGATCAGGCTGCTGTGCTT	1200
Db	121926	AATCACATCAAGACCGCTGACTGGACACTTGCCTATGATGAAGATCAGGCTGCTGTGCTT	121985
QY	1201	GAGAGAGACAAACGGATGATATAACCTCGGAATTAGAAGTGAACAGTGAATCGGAAAT	1260
Db	121986	GAGAGAGACAAACGGATGATATAACCTCGGAATTAGAAGTGAACAGTGAATCGGAAAT	122045
QY	1261	GGTGCTTACTTAGATAAATCTTCCAAAGGAGCTTTGAAGAACTGATTGATCTGCTGTAG	1320
Db	122046	GGTGCTTACTTAGATAAATCTTGAAGGAGCTTTGAAGAACTGATTGATCTGCTGTAG	122105
QY	1321	TTAAATGCTCTTTTAAAAACCTTAGAGGAGAAAGAAACCAAATTTATATTCAGTTGTCT	1380
Db	122106	TTAAATGCTCTTTTAAAAACCTTAGAGGAGAAAGAAACCAAATTTATATTCAGTTGTCT	122165
QY	1381	GAGTTTGATAAACAAGAGAGAGCTTACAGAGCATATTAAGAATCTTCAGACTCAACAA	1440
Db	122166	GAGTTTGATAAACAAGAGAGAGCTTACAGAGCATATTAAGAATCTTCAGACTCAACAA	122225
QY	1441	GCATCTTTCAGTCAAGAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAA	1500
Db	122226	GCATCTTTCAGTCAAGAAACACATTTTGAATAAGAGATCAGAGCTTCAACAGAA	122285
QY	1501	CTTAAAGTAAATGACTGAATATATCAAGAAATGAAATGAACTCCACAGGAATTAACA	1560
Db	122286	CTTAAATAAATGACTGAATTTCTATCAAGAAATGAAATGAACTTCACAGGAATTTAACA	122345
QY	1561	GTAGAGGAAAAATTATCGGTTAGAGAAAGAGAGAACTTTCTTAAAGTAGATGAAAAGATC	1620
Db	122346	GTAGAGGAAAAATTACCGAATAGAGAAAGAGAGAGCTTTCTAGAGTGGAGAAAAGCTC	122405
QY	1621	AGCCATTGCCACTGAAGAGCTGGAGACCTTATAGAAGCGAGCCAAAGATCTTTGAAGAGAA	1680
Db	122406	AGCCGTGCCACTGAACAGCTGGAGACCTTATAGAAGCTAGCCAAAGATCTTGAAGAGAA	122465
QY	1681	TTGGAGAGAACTATTATTTCTATCAAGGCGAGATATTTCCTCCATGAGAAAAAGCACAT	1740
Db	122466	TTGGAGAGAACTTGTTTCATTTTTATCAAAAGCAGAGTATTTCCTACAGAAAAGAGAGACAT	122525
QY	1741	GATAATTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTTAAAGGAAAAGAAAT	1800
Db	122526	GATAATTGGTTGGCAGCTCGGACTGCTGAAGAAACCTCAGTGATTTTAAAGAAAGAAAT	122585
QY	1801	GCTCACACAGACAAAAATTAATGAAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAAGAT	1860
Db	122586	GCTCACACACAAAAATTAATGAAAACAGAGTTGAAATTTGAACTTTTGAAGAAAAGAT	122645
QY	1861	CCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCCTCATATGGTCCC	1920
Db	122646	CCTAATGCACTTGATGTTTCAAAATACAGCATTTGSCAGAGAGCATGCCCCGAATGGTCCC	122705
QY	1921	TCACCAATTGGTTGGCCCTCACTGAAAACAGAGCTTTTCTCTCTCTCCACTTGTGTTG	1980
Db	122706	GCACCAATTGGGTCAGCGTTTCATCTGAAACAGAGAGCTTTTCTCTCTCTCCTCAAACTTGTG	122765
QY	1981	GAGGGTCCACTCAGACTCTCACTTTGCTTCACGGGGGAGAGGAGAAAGGCTCACGAGGC	2040

122766	Db	GAGGATCCACCTGGGACCTCTCACCTGTGCTTCGGAGGGAGGAGGAAGGCCCAAGGC	122825
2041	Qy	CCAGGGAATCCTCTGGACCATCAGATTACCAATTGAAAGAGGAGAAATCAAGCTGTGATAGG	2100
122826	Db	CCAGGGAATCCCTTGGACCATCAGATTACCAATTGAAAGAGGAGAAATCAAGCTGTGATAGG	122885
2101	Qy	TTAACCAGATCTCATAGGGTCCCTCTGACACCTGGGTCTCTGTGCACCTCCATGGGACACAG	2160
122886	Db	TTAACCAGATCTCATAGGGTCCCTCTGACACCTGGGTCTCTGTGCACCTCCATGGGACACAG	122945
2161	Qy	GACCTAGGATGATGTTTCTCGCCAGGACAAATCATATCTGATTCAGGCCCTTCTCTCA	2220
122946	Db	GACTGAAGATGATGTTTCTCGCCAGGACAAATCATATCTGATTCAGGCCCTTCTCTCA	123005
2221	Qy	CAAGGCAAGACAGATTTTGTTCATTTCTGTGTAGACTGTCTGGACCAGACAGAACTCAGA	2280
123006	Db	CAAGGCAAGACAGATTTTGTTCATTTCTGTGTAGACTGTCTGGACCAGACAGAACTCAGA	123065
2281	Qy	AGTTTAAATATGCCCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC	2340
123066	Db	AGTTTAAATATGCCCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC	123125
2341	Qy	AGTGAAGATGATACCAAGATGATCTTTGGTAATTTAAATGTGCTGATTCATCTCTCCCT	2400
123126	Db	AGTGAAGATGATACCAAGATGATCTTTGGTAATTTAAATGTGCTGATTCATCTCTCCCT	123185
2401	Qy	GCTGAAATGAAGCCACTGGCCCTGGCTTGTTCCTCCAGCTTGTCTCCAATCAGAGT	2460
123186	Db	GCTGAAATGAAGCCACTGGCCCTGGCTTGTTCCTCCAGCTTGTCTCCAATCAGAGT	123245
2461	Qy	CCATTGTTTCCAGTGGATGCAAGAGGCCCAATCTTGAGAGAGAGACCTCCTTTCCGCCCA	2520
123246	Db	CCATTGTTTCCAGTGGATGCAAGAGGCCCAATCTTGAGAGAGAGACCTCCTTTCCGCCCA	123305
2521	Qy	CCTCTCCAGGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGATTTCCCA	2580
123306	Db	CCTCTCCAGGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGATTTCCCA	123365
2581	Qy	GGTCCACCACTGCTGCCATTTGCAATGAGAAATGTCTATCCACCGAGGGTTTTCTCTCT	2640
123366	Db	GGTCCACCACTGCTGCCATTTGCAATGAGAAATGTCTATCCACCGAGGGTTTTCTCTCT	123425
2641	Qy	TACCTTCCCGAGACCTGGATTTTCCCGCCACCCCAATCTCGAAGTGAAGTGAG	2700
123426	Db	TACCTTCCCGAGACCTGGATTTTCCCGCCACCCCAATCTCGAAGTGAAGTGAG	123483
2701	Qy	TTCCCTCAGGTTGATTTCCACCTTCAATGAGCGCTGCTACTCAACATCCAGAACACAG	2760
123484	Db	TTCCCTCAGGTTGATTTCCCGCTTTAA-----TGCTACTGAAATCCAGACACACA-	123535
2761	Qy	CAGAACTGCACATATTTTGTCTCTTCAAAAGTAAATTTTGACTGATCTCATTTTCA	2820
123536	Db	CAGAACTGCACATATTTTGTCTCTTCAAAAGTAAATTTTGACTGATCTCATTTTCA	123595
2821	Qy	GTTTAAGTAACCTGCTGTACTTAAGTGATTAACATTTTGTCTCAATTAAGCTTAATGGA	2880
123596	Db	GTTTAAGTAACCTGCTGTACTTAAGTGATTAACATTTTGTCTCAATTAAGCTTAATGGA	123655
2881	Qy	ATTATAATTTCCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGATCTTATGAG	2940
123656	Db	ATTATAATTTCCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGATCTTATGAG	123715
2941	Qy	TAAATTAATTCGAAT--TTATTTAGCGGTATTAATCTATTTCAATTTG-----ATTAAATCC	2994
123716	Db	TAAATTAATTCGAATTTTATTTATTTAGATCATTAATCTATTTTAAATTTGTTGAATTAATCC	123775
2995	Qy	ACTATTATATAACAATAGTGGAGTTTATATATGTAATCTTTTCAGGTGGGAGGCTTT	3054
123776	Db	ACTATTATATAACAATAGTGGAGTTTATATATGTAATCTTTTCAGGTGGGAGGCTTT	123835
3055	Qy	AAATTCGAAGTCT--GTGTCCTTTATCCCAAGAACTGTATTTACTGTGTTGTGACAA	3111

Db 123836 AATCTTAAGGTTGGTCTCTCATGCAAGAACTGATTCTACTGTGTTAGATAA 123895

Qy 3112 ATCTGAAGTAACCTTATGCTTAAATAAATAATAGTTGATTTA 3154

Db 123896 ATGTGAAGTAACCTTATGCTTAAATAAATAATAGTTGATTTA 123938

RESULT 9

HSU94780 3676 bp mRNA linear PRI 20-NOV-1997

LOCUS Human meningioma-expressed antigen 6 (MEA6) mRNA, complete cds.

DEFINITION U94780

ACCESSION U94780

VERSION U94780.1 GI:2231998

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3676)

AUTHORS Heckler, D., Brass, N., Fischer, U., Blin, N., Steudel, I., Tureci, O., Rackler, O., Zang, K.D. and Meese, E.

TITLE cDNA cloning and chromosomal mapping of a predicted coiled-coil proline-rich protein immunogenic in meningioma patients

JOURNAL Hum. Mol. Genet. 6 (12), 2031-2041 (1997)

MEDLINE 97472454

PUBMED 9356211

REFERENCE 2 (bases 1 to 3676)

AUTHORS Heckler, D., Brass, N. and Meese, E.U.

TITLE Direct Submission

JOURNAL Submitted (20-MAR-1997) Inst. for Human Genetics, University of Saarland; Medical School, Oskar-Orth-Strasse, Homburg, Saarland 66421, Germany

FEATURES

Location/Qualifiers

1..3676

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="meningioma"

/notes="PCR and in-situ hybridization mapping locate the gene to chromosomes 2,3,6,7,9,13,14"

1..3676

/gene="MEA6"

315..2729

/gene="MEA6"

/note="meningioma-expressed antigen 6; this sequence is similar to MEA1, GenBank Accession Number U73682, but is longer; it may be an alternate splice variant or it may be from a separate gene"

/codon_start=1

/product="MEA6"

/protein_id="AAB86593.1"

/db_xref="GI:2231998"

/translations="MEEPGATPOPLYGLLEELRRVVAALPEGMRPDSNLYGFPWELV ICAAVGFFAVLFLWRSFRVSRLYVGREKLLMLSLGLIEPKSLKEKLSLVQKE YEGVESSLKDSFKEATEAQSLATECEKLNRSSELEDEILCLEKELEKESKHS EQELMADISKRIQSLDESKLSOVAEAKMTFOIPWNEERLKIAIKDALNENSL QESKOLLOBAEVKVEQSLNKQVTFEDSKVHAEQVLDNKEHSHKTLTERLLKMD KRAMLEDITDDNLELWNSSENGAYLDNPPKGLKLIHAKLNLSIKLEGERN QIYQSEVKTKEELELTHKLNQTSQSLQSENTHPENKQLOKRLKMTLYGREN EMKLRLKLTVEENRLEKEKLSKVDKISHATELELYTKRKAQLEELERTHSYQ QGISHEKKAHDNLWLAARNALNLDLRKENAHNRKLTETELKFLLEKDPYALDVP NTAIFGRHSYPSPGLGWSSSETRAFSLPPTLLEGLPLTISPLLPFGGGRSGRGNPL DHOITNBRGSSCDRLTDPHRLSDTGLSPWQDRMMFPFGQSYDPSALPPQRO DRFCNSGRILSGPAELFSPNRLSLDKMDGMSMESSRNDTKDILGNLNVPSLSIPA ENETAGPQFVPLPVRGFLPVDAGPFLRGPPPPPPPPGAMFGASRDYPPGDF PGPPAPFANRVYPPRGPPPLPPPPPPPPPPHSEGRSEFPFGLIPSPNEPATEHP EPQOET"

misc_feature 1854..1982

/gene="MEA6"

/note="possible alternatively spliced exon"

3135..3140

/gene="MEA6"

/note="first polyA_signal"

polyA_signal 3640..3645

/gene="MEA6"

/note="second polyA_signal"

ORIGIN

Query Match 87.4%; Score 2764.2; DB 9; Length 3676;

Best Local Similarity 94.9%; Pred. No. 0;

Matches 2921; Conservative 0; Mismatches 128; Indels 30; Gaps 5;

Qy 90 CTAAGGCGGCGGCTCAGACCAAGCGCTGCTCAGGATGTAAAGTGTAAACAAGAGGCC 149

Db 57 CTGTGCGGTCGGGTCGGACCTCGCTCGGATGTAAAGTGTAAACAAGAGGCTC 116

Qy 150 AGGGAGGTGTGGGGGACACATGGGCTGTGAGGCTGTGGGTGCCCGGTTCGCCAG 209

Db 117 GGGATGGG-----CAGGTAGGCTGTGAGGCTGCGGCTGCCCTGTCCCCAG 166

Qy 210 CTCGCCCGCAGCGCTCCACAGTGTCCGCTCCGCTGCTTGTTCACGTGCGCATTCGG 269

Db 167 CTCGCCCGCAGCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 222

Qy 270 GTTCCAGACCAAGCTGCGTGTCTCCACCGCTTGTGT-----GGCCAGTGT 319

Db 223 GTTCCGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 282

Qy 320 ACTGCGGTGACCGCCAGACGACCTCGACGCTATGAGGAGCGCTGTGCTACCCCTCAGC 379

Db 283 ACTGCGGCGACCAAGACGACGCTTGGCGCTATGAGGAGCGCTACCCCTCAAC 342

Qy 380 CTTACTGCGGCTGCTGCTGAGGAGCTACGACAGTGTGGCAGCCTACTCTCAGAGTA 439

Db 343 CGTATTTGGGCTGCTGCTGAGGAGCTACGACGCTGCTGAGGAGCTGCTGCTGAGGTA 402

Qy 440 TGAGACGAGTACAGATCTTATGCTTTCATCGGAACCTGCTGCTGCTGCTGCTGCTGCT 499

Db 403 TGAGACGAGTACAGATCTTATGCTTTCATCGGAACCTGCTGCTGCTGCTGCTGCTGCT 462

Qy 500 TTGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559

Db 463 TTGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522

Qy 560 TTATCGTGGGAGAGAGCAAAACCTTGGTGCAACGCTTCTGGAACCTTATGCAAGAAAAAT 619

Db 523 TTATCGTGGGAGAGAGCAAAACCTTGGTGCAACGCTTCTGGAACCTTATGCAAGAAAAAT 582

Qy 620 GTAACTACTTGAATAATTTAGCTTATTTAAAGAGTATGAAGGCTATGAAGTAGT 679

Db 583 GTAACTACTTGAATAATTTAGCTTATTTAAAGAGTATGAAGGCTATGAAGTAGT 642

Qy 680 CATCTTTAGAGTGCAGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 739

Db 643 CATCTTTAGAGTGCAGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 699

Qy 740 CAACCTGTGAAGAGCTGAACAGGCTCAATCTGAACCTGAGGATGAATCTCTGCTAG 799

Db 700 CAACCTGTGAAGAGCTGAACAGGCTCAATCTGAACCTGAGGATGAATCTCTGCTAG 759

Qy 800 AAAAGAGCTTAAACAAGAGAAATCTAAACATCTCAACAGATGAATGATGCGGATA 859

Db 760 AAAAGAGCTTAAACAAGAGAAATCTAAACATCTCAACAGATGAATGATGCGGATA 819

Qy 860 TTTCAAAAGTATACAGTCTCTAGAGATAGTCAAAATCCCTCAAAATCAAAATAGCTG 919

Db 820 TTTCAAAAGTATACAGTCTCTAGAGATAGTCAAAATCCCTCAAAATCAAAATAGCTG 879

Qy 920 AAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGACGACGCGCTATAGCAATAA 979

Db 880 AAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGACGACGCGCTATAGCAATAA 939

Qy 980 AAGATGCTTTGAATGAAATTTCTCAACTTCAGACCAAGCATAAAACAGCTTTTTCAGCAAG 1039

Db 940 AAGATGCTTTGAATGAAATTTCTCAACTTCAGACCAAGCATAAAACAGCTTTTTCAGCAAG 996

RESULT 10

RESULT 10

AC076959 162320 bp DNA linear HTG 27-AUG-2000
 LOCUS Homo sapiens chromosome 3 clone RP11-922L8, WORKING DRAFT SEQUENCE,
 15 unordered pieces.
 AC076959
 AC076959.4 GI:99293333
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 162320)
 AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenhorn,K.,
 Bodota,B., Bock,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
 Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
 David,R., Deigado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
 Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
 Forcum-Tansey,J., Frantz,P., Ganes,R., Gorrell,J.H., Gorrell,L.L.,
 Guevara-Warr,H., Harris,K., Hernandez,J., Hodgson,A., Hogue,M.,
 Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
 Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
 Licharge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
 Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
 Morris,S., Nash,S., Nelson,A., Nguyen,K., Nguyen,N., Nguyen,S.,
 Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
 Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
 Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugeng,R.,
 Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
 Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
 Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
 Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 162320)
 Worely,K.C.
 Submitted
 Submitted (01-AUG-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Aug 26, 2000 this sequence version replaced gi:9743382.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: HARY
 Center clone name: RP11-922L8
 ----- Summary Statistics
 Sequencing vector: M13: L08821
 Chemistry: Dye-Primer Bodipy: 25% of reads
 Assembly: Dye-terminator Big Dye: 75% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 142875 bases at least Q40
 Consensus quality: 150829 bases at least Q30
 Consensus quality: 155246 bases at least Q20
 Estimated insert size: 156861; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 3.7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 23707: contig of 23707 bp in length
 * 23708 23807: gap of unknown length
 * 23808 47475: contig of 23668 bp in length

* 47476 47575: gap of unknown length
 * 47576 47575: contig of 17614 bp in length
 * 47576 47575: gap of unknown length
 * 47576 47575: contig of 15527 bp in length
 * 80817 80916: gap of unknown length
 * 80917 80916: contig of 14577 bp in length
 * 95493 95493: gap of unknown length
 * 95494 95493: contig of 13333 bp in length
 * 95594 108926: gap of unknown length
 * 108927 108926: contig of 11052 bp in length
 * 109027 120078: contig of 11052 bp in length
 * 120079 120178: gap of unknown length
 * 120179 120178: contig of 7029 bp in length
 * 127308 136555: contig of 9248 bp in length
 * 136556 136555: gap of unknown length
 * 136556 142801: contig of 6146 bp in length
 * 142802 142801: gap of unknown length
 * 142902 152041: contig of 9140 bp in length
 * 152042 152141: gap of unknown length
 * 152142 152141: contig of 3047 bp in length
 * 152189 155288: gap of unknown length
 * 155289 156884: contig of 3396 bp in length
 * 156885 158785: gap of unknown length
 * 158785 161217: contig of 2433 bp in length
 * 161218 161217: gap of unknown length
 * 161318 162320: contig of 1003 bp in length.
 FEATURES
 Location/Qualifiers
 1..162320
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-922L8"
 ORIGIN
 Query Match 86.9%; Score 2749.8; DB 2; Length 162320;
 Best Local Similarity 93.9%; Pred. No. 0;
 Matches 2972; Conservative 0; Mismatches 174; Indels 19; Gaps 10;
 QY 1 AGCTTATCATGATGAGAGACAGCAATTTAAAGTCTATTCCACCAAAATAAAGCTT 60
 DB 17562 AGCTTATCATGATGAGAGACAGCAATTTAAAGTCTATTCCACCAAAATAAAGCTT 17621
 QY 61 CAAATATGATGATGAGAACTGCCAGAACTAAGCGCGCGCGGCTCAGACGAGCGTGC 120
 DB 17622 CAAATACGTGATGAGAACTGCCAGAACTAAGCGCGCGGCTCAGACGAGCGTGC 17681
 QY 121 CTCAGATGATAAGTGTAAACAGAGGGCCAGGGAGGTGGTGGGGGACAAATGGGCTG 180
 DB 17682 CTCAGATGATAAGTGTAAACAGAGGGCCAGGGAGGTGGTGGGGGACAAATGGGCTG 17741
 QY 181 TGAGGCTGTGGGTGCGCGGTTCCTCCAGCTCCCCCGCAGCCCGCTCCACAGTGGTCCG 240
 DB 17742 TGAGGCTGTGGGTGCGCGGTTCCTCCAGCTCCCCCGCAGCCCGCTCCACAGTGGTCCG 17801
 QY 241 CTCGGTGTGGTGTCACTGCGGCATTCGGGTTCAGACCCCAAGGCTGGGTCTCCACC 300
 DB 17802 CTCGGTGTGGTGTCACTGCGGCATTCGGGTTCAGACCCCAAGGCTGGGTCTCCACC 17861
 QY 301 GCTTGTGTGGCGCAGTGTACTGCGGTGACCGCAGAGCAGCGCTCCAGCTATGGAGGAG 360
 DB 17862 GCTTGTGTGGCGCAGTGTACTGCGGTGACCGCAGAGCAGCGCTCCAGCTATGGAGGAG 17921
 QY 361 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGTCTGGAGAGCTACGAGAGTTGTG 420
 DB 17922 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGTCTGGAGAGCTACGAGAGTTGTG 17981
 QY 421 GCAGCACTACCTGAGAGTATGAGCCAGATGAGATCCCTTATGTTTCCATCGGAACCTG 480
 DB 17982 GCAGCACTACCTGAGAGTATGAGCCAGATGAGATCCCTTATGTTTCCATCGGAACCTG 18041
 QY 481 GTGGTATGTGCACTGTATTATGAGATTTTGTGTTTCTCTTTTGTGGAGAGTTT 540

Db 18042 GTGGTATGTGCGAGCTGTTATTGGATTTTGTGTTCTCTCTTTTGTGCGAGAAGTTT 18101
QY 541 AGATCGGTTAGAGTTCGGCTTTACGTGGAGAGAGAGCAAAACCTTGTGTCACGCTTCT 600
Db 18102 AGATCGGTTAGAGTTCGGCTTTACGTGGAGAGAGAGCAAAACCTTGTGTCACGCTTCT 18161
QY 601 GGACTAATTTGAAGAAAAATGTAACCTACTTGAAGAAATTTAGCTTTAATCAAAAAGAGTAT 660
Db 18162 GGACTAATTTGAAGAAAAATGTAACCTACTTGAAGAAATTTAGCTTTAATCAAAAAGAGTAT 18221
QY 661 GAAGGCTATGAGTATGAGTATCTTTAGAGGATGCCAGCTTTGAGAGGAGGCGAGCAAA 720
Db 18222 GAAGGCTATGAGTATGAGTATCTTTAGAGGATGCCAGCTTTGAGAGGAGGCGAGCAAA 18281
QY 721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTTGAACAGGTCCAAATCTTGAACCTTGAG 780
Db 18282 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTTGAACAGGTCCAAATCTTGAACCTTGAG 18341
QY 781 GATGAATCTCTGTCTAGAAAAAGACTTAAACAAAGAGAAATCTTAAACATTTCTCAACAA 840
Db 18342 GATGAATCTCTGTCTAGAAAAAGACTTAAACAAAGAGAAATCTTAAACATTTCTCAACAA 18401
QY 841 GATGAATCTGAGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 900
Db 18402 GATGAATCTGAGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 18461
QY 901 CTCAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 18462 CTCAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 18521
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAGCCAT 1020
Db 18522 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAGCCAT 18581
QY 1021 AAACAGCTTTTTCAGCAAGAAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTAAATAA 1080
Db 18582 AAACAGCTTTTTCAGCAAGAAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTAAATAA 18641
QY 1081 CAGAAATTAACATTTGACGACTCCAAAGTACAGCAGACAGTGTCTGAATGATGAAGAA 1140
Db 18642 CAGAAATTAACATTTGACGACTCCAAAGTACAGCAGACAGTGTCTGAATGATGAAGAA 18701
QY 1141 AATCACATCAAGACCCCTGACTGACACTTGCCTAATGATGAAGAAAGTCAAGCTGCTGTCT 1200
Db 18702 AATCACATCAAGACCCCTGACTGACACTTGCCTAATGATGAAGAAAGTCAAGCTGCTGTCT 18761
QY 1201 GAAGAGACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 18762 GAAGAGACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18821
QY 1261 GGTGCTTACTTATGATTAATCTCTCAAAAGGAGCTTTGAAGAAAGTCAATGATGATGATGAT 1320
Db 18822 GGTGCTTACTTATGATTAATCTCTCAAAAGGAGCTTTGAAGAAAGTCAATGATGATGATGAT 18881
QY 1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAGAAAGCAAAATTTATATTAGTGTCT 1380
Db 18882 TTAATGCTTCTTTAAACCTTTAGAGGAGAGAAAGCAAAATTTATATTAGTGTCT 18941
QY 1381 GAAGTTGATAAAACAAAGGAAGAGCTTACAGAGCATTTTAAATCTTTCAGACTCAACAA 1440
Db 18942 GAAGTTGATAAAACAAAGGAAGAGCTTACAGAGCATTTTAAATCTTTCAGACTCAACAA 19001
QY 1441 GCATCTTTGAGTCAAGAAACACACATTTTGAATGAGAAATCAGAACTTCAACAGAA 1500
Db 19002 GCATCTTTGAGTCAAGAAACACATTTTGAATGAGAAATCAGAACTTCAACAGAA 19061
QY 1501 CTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 19062 CTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 19121
QY 1561 GTAGAGAAATTTATCGGTTAGAGAAAGAGAGAAATCTTCTTAAAGTATGATGATGATGAT 1620
Db 19122 GTAGAGAAATTTATCGGTTAGAGAAAGAGAGAGCTTTCTAGAGTGAAGAGAGCTC 19181

QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
Db 19182 AGCCGTGCCACTGAAGAGCTGGAGACCTATAGAAAAGCTAGCCAAAGATCTTTGAAGAGAA 19241
QY 1681 TTGAGAGAACTATTCAATCTTTATCAAGGCGAGATTTATTTCCATGAGAAAAAGACAT 1740
Db 19242 TTGAGAGAACTATTCAATCTTTATCAAAAGCAGGTTATTTCTTACGAGAAAAAGAGACAT 19301
QY 1741 GATAATTTGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTTAAGGAAAGAAAT 1800
Db 19302 GATAATTTGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTTAAGGAAAGAAAT 19361
QY 1801 GCTCACAACAGACAAAAATTAATGAACAGAGCTTTAAATTTGAATTTTGAAGAAAGAT 1860
Db 19362 GCTCACAACAGACAAAAATTAATGAACAGAGCTTTAAATTTGAATTTTGAAGAAAGAT 19421
QY 1861 CTTTATGCTGCTGATGCTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCTCC 1920
Db 19422 CTTTATGCTGCTGATGCTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTAG--TCC 19479
QY 1921 TCACCAATGGGTTGGCTTCTCATCTGAAAACAAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Db 19480 TCACCAATGGGTTGGCTTCTCATCTGAAAACAAGAGCTTTTCTCTCTCTCAACTTTGTTG 19539
QY 1981 GAGGTTCCACTCAGACTCTCACCCTTGTCTTCCAGGGGAGGAGAGAGGCTCACGAGC 2040
Db 19540 GAGGAT--CACTCAGACTCTCACCCTGCTTCCAGGGGAGGAGAGAGGCTCCAGAGC 19598
QY 2041 CCA--GGGAATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAG 2099
Db 19599 CCAGGGGATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAG 19658
QY 2100 GTTAAACGATCTCTAGGGCTCTCTGACACTGGGTCTCTGTACCTCCATCGGAGCA 2159
Db 19659 GTTAAATGATCTCTCAGAGGCTCTCTGACACTGGGTCTCTGTACCTCCATCGGAGCA 19717
QY 2160 GGAACGTTAGGATGATGTTTCTCTCCGAGAGCAATCATATCTCTGATTCAGCCCTCTCTCC 2219
Db 19718 GGAACGTTAGGATGATGTTTCTCTCCGAGAGCAATCATATCTCTGATTCAGCCCTCTCTCC 19777
QY 2220 ACAAAGCAGACAGATTTTGTCTTAATTTCTGTAGACTGTCTGTGAGAGAGAGAGAGAGAGAG 2279
Db 19778 ACAAAGCAGACAGATTTTGTCTTAATTTCTGTAGACTGTCTGTGAGAGAGAGAGAGAGAGAG 19837
QY 2280 AAGTTTAAATATGCTTCTTTTGGATAAATGGAATGGGTCAATGCTTTCAGAAATGGAATC 2339
Db 19838 AAGTTTAAATATGCTTCTTTTGGATAAATGGAATGGGTCAATGCTTTCAGAAATGGAATC 19897
QY 2340 CAGTAGAAATGATACCAAGATGATCTTGATTAATTTAAATGTCCTGATTCATCTCTCC 2399
Db 19898 CAGTAGAAATGATACCAAGATGATCTTGATTAATTTAAATGTCCTGATTCATCTCTCC 19957
QY 2400 TGCTGAAAATGAAGCCACTGGCTTGTCTTCCCTCCACCTCTTGTCTCCATCAGAGG 2459
Db 19958 TGCTGAAAATGAAGCAACTGGCTTGTCTTCCCTCCACCTCTTGTCTCCATCAGAGG 20017
QY 2460 TCCATCTTTCCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2519
Db 20018 TCCATCTTTCCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20076
QY 2520 ACCTCTCCAGAGCCACTGTTTGGAGCTTCTCGAGATTTTTCACCAAGGAGATTTCCC 2579
Db 20077 ACCTCTCCAGAGCCACTGTTTGGAGCTTCTCGAGATTTTTCACCAAGGAGATTTCCC 20136
QY 2580 AGSTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCC 2639
Db 20137 AGSTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCC 20196
QY 2640 TTACCTTTCCCAAGAGCTGATTTTCCCAACCCCACTTCTGAGAGTGAAGTGA 2699
Db 20197 TTACCTTTCCCAAGAGCTGATTTTCCCAACCCCACTTCTGAGAGTGAAGTGAAGTGA 20254

Db	87497	AGGTTTAATTCATGAAGAAGACAGCAATTTTAAAGTGATTTACCAAAAAATACAAC	87433
Qy	61	CAAAATATGTGATGAAAACTGCCAGAACTAAGCGGGCCGGGCTCAGACCCAGCGGTC	120
Db	87437	CAAAATACGTGATGTGAAAACTGCCAGAACTAGGTGGGTTCGGGTTCGACCTCGCGTC	87378
Qy	121	CTCAGGATGTAAAGTGTAACAAGAGGCCAGGGAGTGGTGGGGGACAAATGGGCTG	180
Db	87377	CCCAGGATGTAAAGTGTAAGAAGGGGGCC---GGGAGTGGTGGGGAACAAATGGGCTG	87321
Qy	181	TGAGGCTGTGGGTGGCCCGTTCCTCCAGCTCCCCCGCAGCCCGCTCCACAGTGGTCG	240
Db	87320	TGAGGCTGTGGGTGCCGGTTCCTCCAGCTCCCCCGCAGCCCGCTCCACAGCGTCCG	87261
Qy	241	CTCCGGTGTGTGTACGTGCGCATTCGGGTCCAGACCCAAAGGTGCGTGTTCCTCCACC	300
Db	87260	CTCCGGTGTGGCTGTACGTGCGCATTCGGGTTCAGACCCAAAGGTGCGTGTTCCTCCACC	87201
Qy	301	GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGCAGCAGCGCTCGACGCTATGAGAG	360
Db	87200	TTTTTATTGTGCCAGTGTACTCGGTGACCGCCAGCAGCAGCGCTCGGCGCTATGAGAG	87141
Qy	361	CCTGGTGCTACCCCTCAGCCCTACCTGGGCGTGGTCTCGAGGAGCTACGACAGTTGTG	420
Db	87140	CCCGGTGCTACCCCTCAGCCCACTTGGGCTGGTCTCGAGGAGCTACGACAGTTGTG	87081
Qy	421	GCAGCACTCCTGAGAGTATGAGACCAAGATGAAATCCTTATGTTTCCATCGGAATG	480
Db	87080	GCAGCACTCCTGAGAGTATGAGACCAAGATGCAACCCCTACGGTTTCCATCGGAATG	87021
Qy	481	GTGCTATGTGCAGCTGTATTGGATTTTTGTGTCTCTCTTTTGTGGAGAAATTTT	540
Db	87020	GTGCTATGTGCAGCTGTATTGGATTTTTGTGTCTCTCTTTTGTGGAGAAATTTT	86963
Qy	541	AGATCGGTTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACTTGGTCCAAACGTTTCT	600
Db	86962	AGATCGGTTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACTTGGTCCAAACGTTTCT	86903
Qy	601	GGACTAATTGAAGAAAAATGTAACACTACTTTGAAAAATTTAGCCTTATTCAAAAAGAGTAT	660
Db	86902	GGACTAATTGAAGAAAAATGTAACACTACTTTGAAAAATTTAGCCTTATTCAAAAAGAGTAT	86843
Qy	661	GAAGGCTATGAAGTAG-----	679
Db	86842	GAAGGCTATGAAGTAGCTTTGAGAAGTAAACTACTTGAAGAGTTTAGCCTTATTCAAAA	86783
Qy	680	CATCTTTAGAGGATGCCAGCTTTGAGAGAGGACGACGAGNAGCAGCAAGTTTGGAGG	739
Db	86782	GAGTATGAAGCGATGAAGTAGCTATGAGAAGCGGCGCAGCAGAACGCAAGTTTGGAGG	86723
Qy	740	CAACCTGTAAAAAGCTGAACAGSTCCAATTTCTGAACCTTGAGGATGAAATCTCTCTCTAG	799
Db	86722	CAACCTGTAAAAAGCTGAACAGSTCCAATTTCTGAACCTTGAGGATGAAATCTCTCTCTAG	86663
Qy	800	AAAAGACTTAAACACAGAGAAATCTTAACATTTCTCAACAGATGAATTTGATGGCGGATA	859
Db	86662	AAAAGACTTAAAGAGAGAGAAATCTTAACATTTCTCAACAGATGAATTTGATGGCGGATA	86603
Qy	860	TTTTCAAAAAGTATACAGTCTCTAGAAGATGAGTCAAAAATCCCTCAAAATCACAATAGCTG	919
Db	86602	TTTTCAAAAAGTATACAGTCTCTAGAAGAGAGTCAAAAATCCCTCAAAATCACAATAGCTG	86543
Qy	920	AAGCCAAAATCATCTGCAGACATTTAAATATGATGAGAAACGACGGCTATAGCAATAA	979
Db	86542	AAGCCAAAATCATCTGCAGATATTTAAATATGATGAGAAACGACGGCTATAGCAATAA	86483
Qy	980	AAGATGCTTTCAATGTAAATTTCTCACTTCAGACAAGCCATTAACAGCTTTTTTCAGCAAG	1039
Db	86482	AAGATGCTTTGAATGTAAATTTCTCACTTCAGGAAAGCCAGAAACAGCTTTT---GCAG	86426
Qy	1040	AAGCTGAAGTATGAAGAGGAGAGTGAAGTGAATTAATAACACAGAAAAATAACATTGAAG	1099
Db	86425	AAGCTGAAGTATGAAGAGGAGAGTGAAGTGAATTAATAACACAGAAAAATAACATTGAAG	86366

2190	CTCCGCCAGGACAAATCATATCCTGATTCAGCGCTTCCTCCACAAAGGCAAGACAGATTTT	2239
85286	CTCCACCAAGGACAAATCATATCCTGATTCAGCTTCCTCCCAAAGGAAACACAGATTTT	85227
2240	GTTCCTAAATCTCGGTAGACTGCTGCACAGCAGAACTCAGAAGTTTTTAATATGCTCTCTT	2299
85226	ATTCTAAATCTGAAAGACTGCTGCACAGCAGAAACCCAGAAGTTTTTAAATGACTTCTT	85167
2300	TGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAAATCCAGTAGAAATGATACCAAG	2359
85166	TGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAAATCCAGTAGAAATGATGCCAAG	85107
2360	ATGATCTTTGGTAATTTAAATGTGCCGTGATTCATCTCTCCCTGCTGAAATGAAGCCACTG	2419
85106	ATGATCTTTGGTAATTTAAATGTGCCGTGATTCATCTCTCCCTGCTGAAATGAAGCAACTG	85047
2420	GCCCTGCGCTTGTTCCTCCACTGCTCCAAATCAGAGGTCATTGTTTCCAGTGGATG	2479
85046	GCCCTGCGCTTGTTCCTCCACTGCTCCAAATCAGAGGTCATTGTTTCCAGTGGATA	84987
2480	CAAGAGGCCCATCTTTTGAGAAGAGACCTCTCTTCCGCCCACTCTCCAGAGGCCATGT	2539
84986	CAAGGGGCCGTTTCATGAGAAGAGATCTCTTTTCCGCCCACTCTTCAGAACCACTGT	84927
2540	TTGGAGCTTTCGAGATTTATTTTCCAAAGGATTTCCAGGTCACCACTGCTCCAT	2599
84926	TTGGAGCTTTCGAGGTTATTTTCCAAAGGATTTCCAGGTCACCACTGCTCCAT	84867
2600	TTGCAATCAGAAAAATCTCTATCCACGAGGGTTTTCTTCCTTACCTTCCGCCCAAGACTG	2659
84856	TTGCAATCAGAAAAATCTATCCACCAAGGGTTTTACCTCTTACCTTCATCCGAGACCTG	84807
2660	GATTTTTTCCCCCACCCCACTTCTGAAGGTAGAGTGAGTTCGCCCTCAGGTTTGATTC	2719
84806	GATTTTTT-ACCCCAACCCCACTTCTGAAGGTAGAGGCGAGTTCCCTTCAGGATTGATTC	84749
2720	CACCTTCAAATGAGCCTGCTACTGAACATCCAGAACCAAGCAAGAAACCTGCAAAATTT	2779
84748	CGCCITCAA-----TGCTAGAACATCCAGGACCAAGCAAGAACCTTGACATATT	84696
2780	TTTGCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTTCAGTTTAAGTAATCTGCTGTTA	2839
84695	GTGCTCTCTTCAAAAGTAATTTTGACGGATCTCATTTTTCAGTTTAAGTAATCTGCTATTA	84636
2840	CT---TAAGTGATTACACTTTTGCTCAAAATGGAAGCTTAATGGAATTTAAATTTCTCAGG	2895
84635	CTTAGTAGAATGATTGCACTTTTGCTCAAAATGGAAGTTTAATGGAATTTAGTTCTCAGG	84576
2896	ATAGTATTTTGTAAATAAGATGATTTAAATATGAAATCTTATGAGTAATAATTTTCAATT	2955
84575	ATAGTATTTTGTAAATAAGATGCTTTAAATATGAAATCTTATGAGTAATAATTTTCAATT	84516
2956	---TTATTTTAGCGGTATAACTATTTCATTTCAATTG---ATTAAATCCACTATTATATAACA	3009
84515	TATTATTCTAGATCGTATTAATTTTATTTTGGTGAATTAATCCACTGTTATAGAAACA	84456
3010	ATAGTGGAGTTTTATATATGATTAATCTTTTCAGTGGGAGGCTTTAAATTTCTGAAGTCT-	3068
84455	ATATTGGGAGTTTTATATATGATTAATCTTTTCAGGTTGGGAGGCTTTAAATTTCTAAAGTTG	84396
3069	--GTGCTTTATGCCAAGAACTGATTTTACGTGTTGTTGGAACAATCTGAAAGTAACCTT	3126
84395	TGCTGCTCTCATGCCAAGAACTGATTTTACGTGTTGTTGGAATAATCTGAAAGTAACCTT	84336
3127	TATGCTTAAATAAATATATAGTTGATTTA	3154
84335	TATGCTTAAATTAATAAACTTTAGTTTGA	84308

RESULT 12					
HSM807156					
LOCUS	HSM807156	2931 bp	mRNA	linear	PRI 28-AUG-2003
DEFINITION	Hom sapiens mRNA; cDNA DKF2p686F08252 (from clone DKF2p686F08252);				

complete cds.
BX640994 GI:34365350
BX640994.1
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2931)
Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
The German Human cDNA Consortium
Direct Submission
Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686F08252) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..2931
/locus_tag="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="14q21.1"
/clone="DKFZp686F08252"
/tissue_type="human salivary gland"
/clone_lib="686 (synonym: hicc3). Vector pSPORT1_sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1..2931
/gene="DKFZp686F08252"
310..2484
/gene="DKFZp686F08252"
/note="meningiooma-expressed antigen 6 (MEA6), differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45997.1"
/db_xref="GI:34365351"
/translation="MLSGLIEEMKLLKFLSVQKEYEYEVESLLKQASFEKATEA LKSOVAEKMPFKIQMNEERLKAIKDALNENSOIQSQKOLLQEAQVWQKQVSELN QKVTPEISKYGAQVLQVNEERKSHIKTLFERLLKMKDMAALGEDITDDNLLEWNS SEANGIYDNPFGAKLLIHAKUNASLKTLEGERNQYIQISVDVKTEELTEHIKN LQTKQASLQSENTHPENSKQLQKLVTELYQENMKLHRKIVTEENRLEKGEKQ SKVDKISGHATEELTEYKRAKDLDEELTEIHSYQGGIISHEKAKHDNWLAAERAE NLNDKXENANNRQKLTETKLEKEDPYALDVENTPAFGREHPGSPGLGWSPSE TRAFUSPTLLEGLPLSLPLGGGGRSGRGPNLDHOITNERGESCDRLTDPHRA PSDTGSLSPPWDQDRMMFPFGQSYPDSPALPQQRDFCSNGRLSGPAELRSFNMP LDKMDGSMFSEMRSDTKDDLGNLNVYDSSLPAENEATQGVFPVPLAPIRGLF PVDKAGPIRKRGPPFPFPFGMGAFASRDYFFPRDFPGPPPPAFPMRVNYPYPRGPPY LPFRPGFFPFPFSGRSEFPGLPIPPSPNEPATEHPPEQET"
2890..2895
/gene="DKFZp686F08252"
2910
/gene="DKFZp686F08252"
polysignal
polysite
ORIGIN
Query Match 82.0%; Score 2593.8; DB 9; Length 2931;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 2671; Conservative 0; Mismatches 67; Indels 7; Gaps 3;
QY 415 GTTGGGAGCACTACCTGAGAGTATGAGACGATGAGACATCCCTATGCTTTCCATCG 474
Db 132 GTTGGGAGCACTGCTGAGAGTATGAGACGATGAGACATCCCTATGCTTTCCATCG 191
QY 475 GAACGTGGTATGTCAGCTGTTATGGAATTTTGTGTTCTC-CTTTTTTGTGGAG 533

192 GAAATTGGTGATATGTGCACTGTTGTGGATTTTGTGCTCTTTTGTGAG 251
534 AAGTTTATAGTCGGTTAGAGTCGGCTTTTACGTGGGAAGAGCAAAAACCTTGTGCAAC 593
252 AAGTTTATAGTCGGTTAGAGTCGGCTTTTATGTGGGACGAGAAAAAGCTTGTCTAAT 311
594 GCTTTCTGACTAAATTGAAGAAAAATGTAACCTACTTGAAAAATTTAGCCTTATTCAAA 653
312 GCTTTCTGACTAAATTGAAGAAAAATTAACCTACTTGAAAAATTTAGCCTTGTTCAAA 371
654 AGAGTATGAAGCTATGAAGTATGATGTCCTTTTACAGGATGCGACCTTTGAGAGGAGGC 713
372 AGAGTATGAAGCTATGAAGTATGATGTCCTTTTAAAGATGCGACCTTTGAGAGGAGGC 431
714 AGCAGAGAGACGAGGAGTTTGGAGCACTGTCGAAGCTGGAAGCTGACAGCTCCATTCGA 773
432 AAC---AGAGACAAAGTTTGGAGCACTTGTGAAGCTGGAAGCTGGAAGCTCCATTCGA 488
774 ACTTGAGGATGAATCCTCTGTCTAGAAAAAGACTTTAAAAACAGAGAAATCTTAAACATTC 833
489 ACTTGAGGATGAATCCTCTGTCTAGAAAAAGACTTTAAAAAGAGAGAAATCCAAACATTC 548
834 TCAGACAGATGATGATGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGATGTC 893
549 TGAACAGATGATGATGCGGATATTTCAAAAAGGATACAGTCTCTAGAGATGATGTC 608
894 AAAATCCCTCAATCACAATAGCTGAGCCAAAATCATCTGCAAGACATTTTAAATGAG 953
609 AAAATCCCTCAATCACAATAGCTGAGCCAAAATGACCTTCAAGATATTTCAATGAA 668
954 TGAAGACAGCGGCTATAGCAATAAAGATGCTTTGAATGAATAATCTCAACTTCAGAC 1013
669 TGAAGACAGCTGAGATAGCAATAAAGATGCTTTGAATGAATAATCTCAACTTCAGGA 728
1014 AAGCCATAAACAGCTTTTTCAGCAAGAGCTGAAAGTATGAAAGAGAGAGTGAAGTGA 1073
729 AAGCCAGAACAGCTTTT---GCAAGAGCTGAGTATGAAAGACAGTGAAGTGA 785
1074 TAATAACAGAAATTAACATTTGAAGACTCCAAAGTACAGCAGACAAAGTCTGAATGA 1133
786 TAATAACAGAAATTAACATTTGAAGACTCCAAAGTACATGAGCAAGAGTCTTAAATGA 845
1134 TAAAGAAATACATCAAGACCTGACTGGAACACTTGCCAAATGATGAAGATCAGGCTGC 1193
846 TAAAGAAATCACAATCAAGACTGACTGAAACGCTTTTAAAGATGAAGATTTGGCTGC 905
1194 TGTGCTTGAAGAGACACAAACGATGATGATTAACCTTGAATTTAGAGTGAACAGTGAATC 1253
906 TATGCTTGGAGAGACATAACGATGATGATTAACCTTGAATTTAGAAATGAACAGTGAATC 965
1254 GGAATAATGCTTACTTATGATAATCCTCCAAAGAGCTTTTGAAGAACTGATTCATGC 1313
966 GGAATAATGCTTACTTATGATAATCCTCCAAAGAGCTTTTGAAGAACTGATTCATGC 1025
1314 TGCTAAGTAAATGCTTCTTTTAAACCTTTAGAGGAGAGAAACCAAAATTTATTTCA 1373
1026 TGCTAAGTAAATGCTTCTTTTAAACCTTTAGAGGAGAGAAACCAAAATTTATTTCA 1085
1374 GTTGTCTGAAGTGTATTAACAAAGAGAGCTTACAGAGCATATTTAAATACTTCAGAC 1433
1086 GTTGTCTGAAGTGTATTAACAAAGAGAGCTTACAGAGCATATTTAAATACTTCAGAC 1145
1434 TCACCAAGCATCTTTGAGTCAAGAAACACACATTTTGAATAATGAGAACTTCA 1493
1146 TCACCAAGCATCTTTGAGTCAAGAAACACACATTTTGAATAATGAGAACTTCA 1205
1494 ACAGAACTTAAAGTAAATGCTGAAATATATCAAGAAATGAAATGAAACTCCAGAGAA 1553
1206 ACAGAACTTAAAGTAAATGCTGAAATATATCAAGAAATGAAATGAAACTCCAGAGAA 1265
1554 ATTAACAGTATAGAGAAATATTCGGTTAGAGAAAGAGAACTTTCTTAAAGTATGATGA 1613

1266 ATTAACAGTATAGAGAAATTTATCGTTAGAGAAAGAGAGAACTTTCTTAAAGTATGTA 1325
1614 AAGATCAGCATGCGCACTGAAGAGCTGGAGACTTATAGAAACGAGCCAAAGATCTTGA 1673
1326 AAGATCAGCATGCGCACTGAAGAGCTGGAGACTTATAGAAACGAGCCAAAGATCTTGA 1385
1674 AGAAGAAATGGAGAGAACTATTCAATCTTATCAAGGCGCAGATTTATTTCCCATGAGAAAA 1733
1386 AGAAGAAATGGAGAGAACTATTCAATCTTATCAAGGCGCAGATTTATTTCCCATGAGAAAA 1445
1734 AGCAGATGATTAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1793
1446 AGCAGATGATTAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1505
1794 AGAATAATGCTCAACAACAGACAAAAATAAATCTGAAACAGAGCTTAAATTTGAACTTTTGA 1853
1506 AGAATAATGCTCAACAACAGACAAAAATAAATCTGAAACAGAGCTTAAATTTGAACTTTTGA 1565
1854 AAAAGATCTTATGCACTGATGCTTCCAAATACAGATTTGGCAGAGAGATTTCCCATTA 1913
1566 AAAAGATCTTATGCACTGATGCTTCCAAATACAGATTTGGCAGAGAGATTTCCCATTA 1625
1914 TGGTCCCTCACCATTGGGTTGGCTTTCATCTGAAACAGAGCTTTTCTCTCTCTCCAAC 1973
1626 TGGTCCCTCACCATTGGGTTGGCTTTCATCTGAAACAGAGCTTTTCTCTCTCTCCAAC 1685
1974 TTTGTTGGAGGTCCTCCTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTC 2033
1686 TTTGTTGGAGGTCCTCCTCAGACTCTCCTCAGCTCTCCTTCCAGGGGAGAGAGAGGCTC 1745
2034 ACAGAGCCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAAGAGAGAAATCAAGCTG 2093
1746 ACAGAGCCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAAGAGAGAAATCAAGCTG 1805
2094 TGATAGTTTAAACCGATCCTCATAGGCTCCTCTGACACTGGGTCTCTGTCACTCCATG 2153
1806 TGATAGTTTAAACCGATCCTCATAGGCTCCTCTGACACTGGGTCTCTGTCACTCCATG 1865
2154 GGACCGAGCCGTAGAGATGATTTCTCTCGCAGAGCAATCATATCTCTGATTCAGCCCT 2213
1866 GGACCGAGCCGTAGAGATGATTTCTCTCGCAGAGCAATCATATCTCTGATTCAGCCCT 1925
2214 TCCTCCACAAAGGCAAGACAGATTTTGTCTAATTTGTTAGTCTGTCTGGACAGAGAGA 2273
1926 TCCTCCACAAAGGCAAGACAGATTTTGTCTAATTTGTTAGTCTGTCTGGACAGAGAGA 1985
2274 ACTCAGAGATTTAATATGCTCTTTTGGATAAATGGATGGGTCAATGCTCTTCAAGAAAT 2333
1986 ACTCAGAGATTTAATATGCTCTTTTGGATAAATGGATGGGTCAATGCTCTTCAAGAAAT 2045
2334 GGAATCCAGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATC 2393
2046 GGAATCCAGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATC 2105
2394 TCCTCCCTGCTGAAATGAAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTCTGCTCCAAT 2453
2106 TCCTCCCTGCTGAAATGAAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTCTGCTCCAAT 2165
2454 CAGAGGTCCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGAGACCTCTT 2513
2166 CAGAGGTCCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGAGACCTCTT 2225
2514 CCCCCACCTCTCCAGAGCCATTTGTTGAGCTTCTCGAGATTTATTTCCACCAAGGGA 2573
2226 CCCCCACCTCTCCAGAGCCATTTGTTGAGCTTCTCGAGATTTATTTCCACCAAGGGA 2285
2574 TTTCCAGGTCCACCACTCTCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTT 2633
2286 TTTCCAGGTCCACCACTCTCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTT 2345
2634 TCCTCTTTA CTTTCCCCCAAGACCTGATTTTTCCTCCCCCAAGGCTTCTGAGGTTAG 2693
2346 TCCTCTTTA CTTTCCCCCAAGACCTGATTTTTCCTCCCCCAAGGCTTCTGAGGTTAG 2405

2694 AAGTGAGTTCCTCCAGCTTTGATTCACACCTTCAATAGAGCCTGCTACTGACATCCAGA 2753
 Db AAGTGAGTTCCTCCAGCTTTGATTCACACCTTCAATAGAGCCTGCTACTGACATCCAGA 2465
 QY ACCACAGCAAGAACCTGACATATTTTGGCTCTTTCARAGTAAATTTGACTGATCTC 2813
 Db ACCACAGCAAGAACCTGACATATTTTGGCTCTTTCARAGTAAATTTGACTGATCTC 2525
 QY ATTTCAGTTCCTTAAAGTAACTGCTTACTTAAAGTAACTTAAAGTAACTTAAAGT 2873
 Db ATTTCAGTTCCTTAAAGTAACTGCTTACTTAAAGTAACTTAAAGTAACTTAAAGT 2585
 QY TAATGAATATTAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGATC 2933
 Db TAATGAATATTAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGATC 2645
 QY TTATGAGTAAATTTTAAATTTTATTTTATAGCGGTATTAATTTCAATTTGATTAATC 2993
 Db TTATGAGTAAATTTTAAATTTTATTTTATAGCGGTATTAATTTCAATTTGATTAATC 2705
 QY CACTATTATTAATTAAGTAACTGCTTACTTAAAGTAACTTAAAGTAACTTAAAGT 3053
 Db CACTATTATTAATTAAGTAACTGCTTACTTAAAGTAACTTAAAGTAACTTAAAGT 2765
 QY TAAATCTGAGTCTGCTTTTATGCGAAGAACTGATTTACTGTGTTGTGGACAAAT 3113
 Db TAAATCTGAGTCTGCTTTTATGCGAAGAACTGATTTACTGTGTTGTGGACAAAT 2825
 QY GTGAAAGTAACTTTATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 3158
 Db GTGAAAGTAACTTTATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 2870

RESULT 13
 HSM808160 2957 bp mRNA linear PRI 30-AUG-2003
 LOCUS Homo sapiens mRNA; cDNA DKF2p686M2159 (from clone DKF2p686M2159).
 DEFINITION BX648014
 ACCESSION BX648014
 VERSION BX648014.1 GI:34367173
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2957)
 AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
 Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
 Wiemann,S.
 The German Human cDNA Consortium
 Direct Submission
 Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKF2p686M2159) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.
 Location/Qualifiers
 1..2957
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKF2p686M2159"
 /tissue_type="human endometrium"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
 DH10B; sites SfiI + SfiIB"
 /dev_stage="adult"

polyA_signal 2846..2851
 polyA_site 2864
 ORIGIN
 Query Match 81.6%; Score 2583; DB 9; Length 2957;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 2673; Conservative 0; Mismatches 70; Indels 8; Gaps 4;
 QY 415 GTTGTGCGAGCACTACCTGAGAGTATGAGACCGAGATGAGAACTCTTATGCTTTCCATCG 474
 Db 132 GTTGTGCGAGCACTGCGCTGAAGGTATGAGACCGAGATCTAATCTTTATGCTTTCCATGG 191
 QY 475 GAACTGTGGTATGTCAGAGCTGTTATGGAATTTTGTGCTCTC-CTTTTGTGCGAG 533
 Db 192 GAATTTGTGATGCTGAGCTGTTGTTGGAATTTTGTGCTCTCTTTTGTGCGAG 251
 QY 534 AAGTTTATAGTCGGTTAGGAGTCGGCTTTACGTCGGGAAGAGACAAAACCTTGTGCGAAC 593
 Db 252 AAGTTTATAGTCGGTTAGGAGTCGGCTTTATGTGGGACGAGAGAAAAGCTTGTCTTAAT 311
 QY 594 GCTTTCGGACTAATTCGAAGAAAATGTAACACTTACTTTGAAAATTTTACGCTTATTCAAA 653
 Db 312 GCTTTCGGACTAATTCGAAGAAAATGTAACACTTACTTTGAAAATTTTACGCTTATTCAAA 371
 QY 654 AGAGTATGAAGGCTATGAAGTAGAGTCACTTTAGAGGATGCCAGCTTGAAGAGGAGGC 713
 Db 372 AGAGTATGAAGGCTATGAAGTAGAGTCACTTTAAGGATGCCAGCTTGAAGAGGAGGC 431
 QY 714 AGCAGAGACGACGAGTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTCGA 773
 Db 432 AAC---AGAGCACAAAGTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTCGA 488
 QY 774 ACTTGAGGATGAAATCCCTCTCTCTAGAAAAGAGCTTAAAACAGAGAAATCTAAACATTC 833
 Db 489 ACTTGAGGATGAAATCCCTCTCTCTAGAAAAGAGCTTAAAACAGAGAAATCTAAACATTC 548
 QY 834 TCACACAGAGTGAATGATGGCGGATATTTCAAAAGTATACAGTCTCTAGAGATGAGTC 893
 Db 549 TGAACAGAGTGAATGATGGCGGATATTTCAAAAGGATACAGTCTCTAGAGATGAGTC 608
 QY 1014 AAGCCATAAACAGCTTTTTCAGCAAGAGCTGAAAGTATGGAAGGAGAGAGTGAAGTGA 1073
 Db 729 AAGCCAGAAACAGCTTTT---GCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGA 785
 QY 1074 TAATAACAGAAAATAACATTTGAAGACTCCAAAGTACACCCAGAACAGTCTTGAATGA 1133
 Db 786 TAATAACAGAAAATAACATTTGAAGACTCCAAAGTACATCAGAACAGTCTTGAATGA 845
 QY 1134 TAAAGAAATACATCAAGACCCCTGACCTGGACACTTGGCAATGATGAAGATCAGGCTGC 1193
 Db 846 TAAGAAAGTCACTCAAGACTCTGACTGAACGCTTGTAAAGATGAAGATTCGGCTGC 905
 QY 1194 TGTGCTTGAAGAGACACCAAGATGATGATACTCGAAATAGAGTGAAGTGAACAGTGAATC 1253
 Db 906 TATGCTTGGAGAAGACATAACCGATGATCACTTGAATTTAGAAATGAACAGTGAATC 965
 QY 1254 GGAAATGCTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGC 1313
 Db 966 GGAAATGCTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGC 1025
 QY 1314 TGTGAAGTTAAATGCTTCTTTAAAAACCTTAGAAGGAGAGAAACCAAAATTTATATCA 1373
 Db 1026 TGTGAAGTTAAATGCTTCTTTAAAAACCTTAGAAGGAGAGAGAAACCAAAATTTATATCA 1085
 QY 1374 GTTGCTGAGTGTGATAAACAAGAGAGGCTTACAGAGCATATTTAAAAATCTTCAGAC 1433

1086 GTTGTCTGAAGTTGATATAAAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGAC 1145
 1434 TCACCAAGCATCTTTGCAAGTCAGAAAACACACATTTTGAAATCAGAAATCAGAACTTCA 1493
 1146 TGAACACAGCATCTTTGCAAGTCAGAAAACACACATTTTGAAATCAGAAATCAGAACTTCA 1205
 1494 ACAGAAAATTAAAGTAATGACTGAATATATCAAGAAAATGAATGAATCAACAGGAA 1553
 1206 ACAGAAAATTAAAGTAATGACTGAATATATCAAGAAAATGAATGAATCAACAGGAA 1265
 1554 ATTAACAGTAGAGGAAAATTTATCGGTTAGAGAAAAGAGAAAATCTTTAAAGTAGATGA 1613
 1266 ATTAACAGTAGAGGAAAATTTATCGGTTAGAGAAAAGAGAAAATCTTTAAAGTAGATGA 1325
 1614 AAAGATCAGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGA 1673
 1326 AAAGATCAGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGA 1385
 1674 AGAAGAAATGGAGAGAACTATTCTTCTATCAAGGGCAGATTTATCCCATGAGAAAAA 1733
 1386 AGAAGAAATGGAGAGAACTATTCTTCTATCAAGGGCAGATTTATCCCATGAGAAAAA 1445
 1734 AGCACAATGATTAATGGTTGGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1793
 1446 AGCACAATGATTAATGGTTGGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1505
 1794 AGAAGATGCTCAACAGACAAAAATTAATCAAGAGAGCTTAAATTTGAATTTTGA 1853
 1506 AGAAGATGCTCAACAGACAAAAATTAATCAAGAGAGCTTAAATTTGAATTTTGA 1565
 1854 AAAAGATCCTTTATGCACTCGATGTTCCAAATACAGATTTGGCAGAGAGCAATCCCCATA 1913
 1566 AAAAGATCCTTTATGCACTCGATGTTCCAAATACAGATTTGGCAGAGAGCAATCCCCATA 1625
 1914 TGGTCCCTCAGCANTGGGTTGGCTTCATCTGAAACAGAGCTTTCTCTCTCCCTCCCAAC 1973
 1626 TGGTCCCTCAGCANTGGGTTGGCTTCATCTGAAACAGAGCTTTCTCTCTCCCTCCCAAC 1685
 1974 TTTGTTGGAGGTCCTCAGACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTC 2033
 1686 TTTGTTGGAGGTCCTCAGACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTC 1745
 2034 ACAGAGCCAGGAAATCTCTGACCATCAGATACCAATGAAGAGGAGATCAGCTG 2093
 1746 ACAGAGCCAGGAAATCTCTGACCATCAGATACCAATGAAGAGGAGATCAGCTG 1805
 2094 TGATAGTTAAACCGATCCTCATAGGCTCCTCTGACACTGGGTCTCTCTCACCTCCATG 2153
 1806 TGATAGTTAAACCGATCCTCATAGGCTCCTCTGACACTGGGTCTCTCTCACCTCCATG 1865
 2154 GGACCCAGGACCGTAGGATGATTTCTCTCGCCAGGACATCATATCTCTGATTCAGCCCT 2213
 1866 GGACCCAGGACCGTAGGATGATTTCTCTCGCCAGGACATCATATCTCTGATTCAGCCCT 1925
 2214 TCCTCCCAAGAGGCAACAGAGATTTGTTCTTAATCTGAGACTGTCTGGACAGCAGAGA 2273
 1926 TCCTCCCAAGAGGCAACAGAGATTTGTTCTTAATCTGAGACTGTCTGGACAGCAGAGA 1985
 2274 ACTCAGAAAGTTTAAATGCTCTCTTTGGATATAAATGATGAGTGGGTCAATGCTTCAGAAAT 2333
 1986 ACTCAGAAAGTTTAAATGCTCTCTTTGGATATAAATGATGAGTGGGTCAATGCTTCAGAAAT 2045
 2334 GGAATCAGTAGAATATCATACCAAGATGATCTTTGTAATTTAAATGTCCTGATTCATC 2393
 2046 GGAATCAGTAGAATATCATACCAAGATGATCTTTGTAATTTAAATGTCCTGATTCATC 2105
 2394 TCTCCCTGCTGAAAATGAAGCAATGCGCCCTGGCTTTGTTTCCCTCCACTCTTCTGCTCAAT 2453
 2106 TCTCCCTGCTGAAAATGAAGCAATGCGCCCTGGCTTTGTTTCCCTCCACTCTTCTGCTCAAT 2165
 2454 CAGAGGTCCTATGTTTCCAGTGGATGCAAGAGGCCCATCTCTGAGAGAGAGACTCTCTTT 2513

2166 CAGAGGTCAATTTGTTTCCAGTGGATGCAAGAGGCCCATCTTCTTGAGAGAGGACCTCTTT 2225
 2514 CCCCCACCTCTCCAGAGGACCATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGGA 2573
 2226 CCCCCACCTCTCCAGAGGACCATGTTTGGAGCTTCTCGAGATTTATTTCCACCAAGGGA 2285
 2574 TTTTCCAGGTCCACACCTGCTCCATTTGCAATGAGAAATGCTTATCCACCCAGGAGGTTT 2633
 2286 TTTTCCAGGTCCACACCTGCTCCATTTGCAATGAGAAATGCTTATCCACCCAGGAGGTTT 2345
 2634 TCTCTCTTACCTTCCCTCCAGGACCTGGATTTTTCCTCCCTCCCTCCCTCCCTCCCTCCCT 2693
 2346 TCTCTCTTACCTTCCCTCCAGGACCTGGATTTTTCCTCCCTCCCTCCCTCCCTCCCTCCCT 2405
 2694 AGGTGAGTTTCCCTCCCTCCAGGACCTGGATTTTTCCTCCCTCCCTCCCTCCCTCCCTCCCT 2753
 2406 AGGTGAGTTTCCCTCCCTCCAGGACCTGGATTTTTCCTCCCTCCCTCCCTCCCTCCCTCCCT 2465
 2754 ACCACAGCAAGAAACCTGACCAATTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2813
 2466 ACCACAGCAAGAAACCTGACCAATTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2524
 2814 ATTTTCCAGTTTAAAGTAACTGCTGTTTAAAGTAACTGCTGTTTAAAGTAACTGCTGTTTAA 2873
 2525 ATTTTCCAGTTTAAAGTAACTGCTGTTTAAAGTAACTGCTGTTTAAAGTAACTGCTGTTTAA 2584
 2874 TAATGGAATTTAATTTCTCAGGATAGTATTTTGAATAAAGATGATTTTAAATATGAAATC 2933
 2585 TAATGGAATTTAATTTCTCAGGATAGTATTTTGAATAAAGATGATTTTAAATATGAAATC 2644
 2934 TTATGAGTAAATTTTCAATTTTATTTTATAGCGGTATTAATTTTCAATTTTCAATTTTCAAT 2993
 2645 TTATGAGTAAATTTTCAATTTTATTTTATAGCGGTATTAATTTTCAATTTTCAATTTTCAAT 2704
 2994 CACTATTATATAAACAATAGTGGAGTTTATATATGTAATATCTTTCAGTGGGAGGCTT 3053
 2705 CACTATTATATAAACAATAGTGGAGTTTATATATGTAATATCTTTCAGTGGGAGGCTT 2764
 3054 TAATTTCTGAGTGTGCTTTTATGTCAGAGACTGTTTATTTACTGTTGTTGTCGACAAAT 3113
 2765 TAATTTCTGAGTGTGCTTTTATGTCAGAGACTGTTTATTTACTGTTGTTGTTGTCGACAAAT 2824
 3114 GTGAAAGTAACTTTATGCTTTAAATAAATATAGTGTGTTTAAATAAATAAATAAATAAATAA 3164
 2825 GTGAAAGTAACTTTATGCTTTAAATAAATATAGTGTGTTTAAATAAATAAATAAATAAATAA 2875

RESULT 14
 HSM803540
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 HOMO SAPIENS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 2890)
 Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Pösch, G., Han, M. and Wiemann, S.
 Direct Submission
 Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFZ686F2453) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.

HSM803540 2890 bp mRNA linear PRI 13-MAY-2003
 Homo sapiens mRNA; cDNA DKFZ686F2453 (from clone DKFZ686F2453).
 AL812233
 AL812233.1 GI:21732779

FEATURES
source

Location/Qualifiers
1..2890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686f2453"
/clone_type="human endometrium"
/clone_lib="686 (synonym: hicc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
2850..2855
polyA_signal
polyA_site
2870

Query Match		80.0%; Score 2531.8; DB 9; Length 2890;
Best Local Similarity		96.3%; Pred. No. 0;
Matches 2644; Conservative		0; Mismatches 72; Indels 29; Gaps 4;
QY	415	GTGTGGCAGCACTACCTGAGAGTATGAGACAGATGAGAACTTATGTGTTCCATCG 474
DB	114	GTGTGGCAGCACTGCTGAAGTATGAGACAGATCTTAATCTTTATGGTTTCCATGG 173
QY	475	GAATCGTGATGTGCAGCTGTTATGGATTTTTTGTGTTCTCTCTTTTTTGTGGAGA 534
DB	174	GAATCGTGATGTGCAGCTGTTATGGATTTTTTGTGTTCTCTCTTTTTTGTGGAGA 233
QY	535	AGTTTATAGTCGGTATGAGTTCGGCTTACGTGGGAGAGAGCAAAACTTGGTCCACG 594
DB	234	AGTTTATAGTC-----GGAGAGAGAGAAAGCTTGGCTCTAATG 271
QY	595	CTTTCTGGACTAATTGAAGAAAAATTAACCTACTTTGAAAAATTTAGCCCTTATCAAAA 654
DB	272	CTTTCTGGACTAATTGAAGAAAAATTAACCTACTTTGAAAAATTTAGCCCTTATCAAAA 331
QY	655	GAGTATGAAGGCTATGAAGTAGTATCTTTAGAGATGCCAGCTTTGAGAGAGGCA 714
DB	332	GAGTATGAAGGCTATGAAGTAGTATCTTTAGAGATGCCAGCTTTGAGAGAGGCA 391
QY	715	GCAGAGAGACACGAAAGTTTGGAGGCAACCTGTGAAAAAGCTGAACAGGTCCTGAA 774
DB	392	AC---AGAGACACAAAGTTTGGAGGCAACCTGTGAAAAAGCTGAACAGGTCCTGAA 448
QY	775	CTTGAGATGAATCTCTCTCTAGAAAAAGACTTAAACAGAGAAATCTAAGCATCT 834
DB	449	CTTGAGATGAATCTCTCTCTAGAAAAAGACTTAAACAGAGAAATCTAAGCATCT 508
QY	835	CAACAGATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGATCA 894
DB	509	GAACAGATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGATCA 568
QY	895	AAATCCCTCAATACAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGAGT 954
DB	569	AAATCCCTCAATACAAATAGCTGAAGCCAAATCATCTGCAAGATATTTCAATGAAT 628
QY	955	GAAGAACGCGGCTATAGCAATAAAGATGCTTTGAATGAAATCTCAACTTCAGACA 1014
DB	629	GAAGAACGCTGAACATAGCAATAAAGATGCTTTGAATGAAATCTCAACTTCAGAAA 688
QY	1015	AGCCATAACAGACTTTTTCAGCAAGAGAGCTGAAGTATGGAAGAGAGAGTGAAGT 1074
DB	689	AGCCAGAAAACGGCTTTT---GCAAGAGCTCAAGTATGGAAGAGAGAGTGAAGT 745
QY	1075	AATAAACAGAAATTAACATTTGAAGCTCCAAAGTACACGAGCAAGTTCTCAATGAT 1134
DB	746	AATAAACAGAAATTAACATTTGAAGCTCCAAAGTATCAATGAGATTTCTAATGAT 805
QY	1135	AAAGAAATACATCAAGACCTGACTGGACACTTGGCAATGATGAAGATCAGGCTGCT 1194
DB	806	AAAGAAATACATCAAGACTCTGACTGAACGCTTTTAAAGATGAAGATTTGGCTGCT 865
QY	1195	GTGCTTGAAGAGACACAAACGATGATGATCACTGGAATTAGAAGTGAACAGTGAATCG 1254
DB	866	ATGCTTGAAGAGACATCAACGATGATGATCACTTGGAAATTTAGAAATGAACAGTGAATCG 925

QY	1255	GAATAATGGTCTTACTTAGATAATCTCCAAAAAGAGAGCTTTGAAGAAACCTGATTCATGCT 1314
DB	926	GAATAATGGTCTTACTTAGATAATCTCCAAAAAGAGAGCTTTGAAGAAACCTGATTCATGCT 985
QY	1315	GCTAAGTTAAATGCTCTTTTAAAAAACCCTTAGAAGGAGAAAGAAACCAAAATTTATATTTCAG 1374
DB	986	GCTAAGTTAAATGCTCTTTTAAAAAACCCTTAGAAGGAGAAAGAAACCAAAATTTATATTTCAG 1045
QY	1375	TTGTCTGAAGTTGATAAAAAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACT 1434
DB	1046	TTGTCTGAAGTTGATAAAAAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACT 1105
QY	1435	CAACAAGCATCTTTCAGTCAGAAAAACACACATTTTGAATAATGAGAAATCAGAAGCTTCAA 1494
DB	1106	GAACAAGCATCTTTCAGTCAGAAAAACACACATTTTGAATAATGAGAAATCAGAAGCTTCAA 1165
QY	1495	CAGAAACTTAAAGTAATGACTGAAATATATCAAGAAAAATGAAATGAAACTCCACAGAAA 1554
DB	1166	CAGAAACTTAAAGTAATGACTGAAATATATCAAGAAAAATGAAATGAAACTCCACAGAAA 1225
QY	1555	TTAACAGTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAA 1614
DB	1226	TTAACAGTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAA 1285
QY	1615	AAGATCAGCCATGCCACTGAAGAGCTGAGACCTTATAGAAAGCGAGCCAAAGATCTTTGAA 1674
DB	1286	AAGATCAGCCATGCCACTGAAGAGCTGAGACCTTATAGAAAGCGAGCCAAAGATCTTTGAA 1345
QY	1675	GAAAGATTGGAGAACTATTCATCTTATCAAGGCGAGATTTATTTCCCATGAGAAAAA 1734
DB	1346	GAAAGATTGGAGAACTATTCATCTTATCAAGGCGAGATTTATTTCCCATGAGAAAAA 1405
QY	1735	GCATCATATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGAAA 1794
DB	1406	GCATCATATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGAAA 1465
QY	1795	GAAAGATTGCTCAACACAGAAATAATTAAGTGAACAGAGCTTAAATTTGAATTTAGAA 1854
DB	1466	GAAAGATTGCTCAACACAGAAATAATTAAGTGAACAGAGCTTAAATTTGAATTTAGAA 1525
QY	1855	AAAGATCCTTATGACCTCGATGTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATAT 1914
DB	1526	AAAGATCCTTATGACCTCGATGTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATAT 1585
QY	1915	GGTCCCTCACATTTGGTGGCCTTCATCTGAAAACAAGAGCTTTTCTCTCTCTCCCAACT 1974
DB	1586	GGTCCCTCACATTTGGTGGCCTTCATCTGAAAACAAGAGCTTTTCTCTCTCTCCCAACT 1645
QY	1975	TTGTTGGAGGTTCCACTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTCA 2034
DB	1646	TTGTTGGAGGTTCCACTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTCA 1705
QY	2035	CGAGCCCGAGGAAATCTCTCGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGT 2094
DB	1706	CGAGCCCGAGGAAATCTCTCGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGT 1765
QY	2095	GATAGTTAAACGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGG 2154
DB	1766	GATAGTTAAACGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGG 1825
QY	2155	GACCAAGACCGTAGATGATGTTTCTCCGCGAGAGCAATCATATCTCTGATTCAGCCCTT 2214
DB	1826	GACCAAGACCGTAGATGATGTTTCTCCGCGAGAGCAATCATATCTCTGATTCAGCCCTT 1885
QY	2215	CTCCACAAAGGCAAGACAGATTTTGTCTTAATCTGCTAGACTGCTCTGGACAGAGAA 2274
DB	1886	CTCCACAAAGGCAAGACAGATTTTGTCTTAATCTGCTAGACTGCTCTGGACAGAGAA 1945
QY	2275	CTCAGAAAGTTTAAATATGCTTTCTTTGGATAAAAATGGATGGGTCAATGCTTCAGAAATG 2334
DB	1946	CTCAGAAAGTTTAAATATGCTTTCTTTGGATAAAAATGGATGGGTCAATGCTTCAGAAATG 2005

2335 GAATCAGTAGAATGATACCAAGATGATCTTGCTGTAATTTAAATGTGCTGATTCATCT 2394
|||||
2394 GAATCAGTAGAATGATACCAAGATGATCTTGCTGTAATTTAAATGTGCTGATTCATCT 2065
|||||
2395 CTCCTGCTGTAATGAAAGCACTGCGCTGCTTGTCTCTCCACCTCTTCTCCCAATC 2454
|||||
2454 CTCCTGCTGTAATGAAAGCACTGCGCTGCTTGTCTCTCCACCTCTTCTCCCAATC 2125
|||||
2455 AGAGTCCATGCTTCCAGTGAATGAGAGCCCATCTTCTGAGAGGACCTCTTCTC 2514
|||||
2514 AGAGTCCATGCTTCCAGTGAATGAGAGCCCATCTTCTGAGAGGACCTCTTCTC 2185
|||||
2515 CCCCCACCTCTCCAGGAGCCATGTTGAGAGCTTCTCGAGATTTATTTCCACCAAGGAT 2574
|||||
2574 CCCCCACCTCTCCAGGAGCCATGTTGAGAGCTTCTCGAGATTTATTTCCACCAAGGAT 2245
|||||
2575 TTCCAGGTCACCACTCTCTCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTT 2634
|||||
2634 TTCCAGGTCACCACTCTCTCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTT 2305
|||||
2635 CTTCTCTTACCTTCCCCCAAGACTGATTTTCCCCCCCCACCCACCATTTCTGAAGGTAGA 2694
|||||
2694 CTTCTCTTACCTTCCCCCAAGACTGATTTTCCCCCCCCACCCACCATTTCTGAAGGTAGA 2365
|||||
2695 AGTGAG-TTCCCTCAGGTTGATTTCCACCTTCAATGAGCTGCTACTGAACATCCAGA 2753
|||||
2753 AGTGAGTTTCCCTCAGGTTGATTTCCACCTTCAATGAGCTGCTACTGAACATCCAGA 2425
|||||
2754 ACCACAGCAAGAACTGACATATTTTTCCTCTCTTCAAAAGTAAATTTGACTGATCTC 2813
|||||
2813 ACCACAGCAAGAACTGACATATTTTTCCTCTCTTCAAAAGTAAATTTGACTGATCTC 2485
|||||
2814 ATTTTCAGTTTAAAGTAACTGCTGTTTAAAGTAACTGATTTTTCCTCAATTTGAAGCT 2873
|||||
2873 ATTTTCAGTTTAAAGTAACTGCTGTTTAAAGTAACTGATTTTTCCTCAATTTGAAGCT 2545
|||||
2874 TAATGAATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTTAAATATCAATC 2933
|||||
2933 TAATGAATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTTAAATATCAATC 2605
|||||
2934 TTATGAGTAAATTTTCAATTTTATGAGGATTAATCAATTTTCAATTTGATTAATC 2993
|||||
2993 TTATGAGTAAATTTTCAATTTTATGAGGATTAATCAATTTTCAATTTGATTAATC 2665
|||||
2994 CACTATTAATAACATAGTGGAGTTTATATATGTAATCTTTTCAGGTGGGAGGCTT 3053
|||||
3053 CACTATTAATAACATAGTGGAGTTTATATATGTAATCTTTTCAGGTGGGAGGCTT 2725
|||||
3054 TAAATTTCTGAGTCTGTGCTTTTATGCCAAGACTGATTTTCTGTTGTTGGACAAT 3113
|||||
3113 TAAATTTCTGAGTCTGTGCTTTTATGCCAAGACTGATTTTCTGTTGTTGGACAAT 2785
|||||
3114 GTCAAGTAACTTTATGCTTAAATTAATTAATGATTTTAAAT 3158
|||||
2786 GTCAAGTAACTTTATGCTTAAATTAATTAATGATTTTAAAT 2830
|||||

RESULT 15
BC064355
LOCUS BC064355
DEFINITION Homo sapiens cDNA clone MGC:75562 IMAGE:4803802, complete cds.
ACCESSION BC064355
VERSION BC064355.1 GI:39963693
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2895)
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Akerman, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Lu, X., Gibbs, R.A.,
Worley, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Villalón, D.K., Metzger, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Kettner, M., Madan, A., Rodriguez, S.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerk, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2895)
Strausberg, R.
Direct Submission
Submitted (15-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, X., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maekari, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripoop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, O.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 134 Row: n Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

Location/Qualifiers
1. 2895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:75562 IMAGE:4803802"
/tissue_type="Skin, normal"
/clone_lib="NCI CGAP_Skn3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
283 2712
/codon_start=1
/product="Unknown (protein for MGC:75562)"
/protein_id="AAH64355.1"
/db_xref="GI:39963693"
translation="MEEPGATPQDLGLLELRVVAALPEGRPDNSLYGFWELV
ICAAVGFVAVLFFLMRFSVRSLYVREKGLALMSGLIEKSLKLPFLVQKE
VEGVEVSSLLKDAEFKEATEQSLEVENQVATCEKLNRSNLEDEILCLELKEE
KSKSEODELMADIKEIQSLEDESKLSQVAEAKMTFKI FOMNERLNIATKDALN
ENSOQSQKQLLOEASVWKEQVSELNKNQVTFEDSKVAEOVLNDKESHIKTLTEL
LKMWDAAALMGEDITDDNLELENSENGALDNPFGALKKLHAAKLNASLKL
EGERNYIQLSEVDKTEKLTETIKNLQEQASQSENFENENQKQKQKVMTE

FEATURES

source

CDS

misc_feature

LYOENMKLHRLKTVENVYRLKEEKLKSKVDEKISHATEELTYRKRANDLEEELERT
IH5YQOITISHEKKAHNDLAANAERNLIRKENAHRNQLLTETELFLELEKXDYP
IADYQATGFRSHSPVGLPGLPSPSETAFSLPTLLLEGRLSPLLPGLGGGRSGR
PGNLDHQTINRGRESSCDRLTDPHRAQDTCGLSPDWDQDRMMPPPPGQYPO5AL
PQORDCFSNGLRSGPAAELSHFNWSESLDKMGSPSESSRRNDTKDLGLNANVPD
SLPAAENATGCGFVPPPLAVYRGVLPFYVDAGPELRCGPPEPPPPGAFGASRDYF
PGDFPGPPPPAPPMANVTTPRGFPPLPFRGFFPPPPHSGRSEFSGSLGPPPSNEP
AIEHPBPQOQT"

1309..1749

/note="APG6; Region: Autophagy protein Apg6. In yeast, 15
Apg proteins coordinate the formation of autophagosomes.
Autophagy is a bulk degradation process induced by
starvation in eukaryotic cells. Apg6/Vps30p has two
distinct functions in the autophagic process, either
associated with the membrane or in a retrieval step of the
carboxypeptidase Y sorting pathway"
/db xref="CDD:pfam04111"

ORIGIN

Query Match	80.0%	Score	2531.4	DB	9	Length	2895
Best Local Similarity	93.9%	Pred. No.	0				
Matches	2702	Conservative	0	Mismatches	136	Indels	39
Gaps	5						
QY	90	CTAAGCGGGCGGGCTCAGACAGCGCTCGCTCAGGATGTAAGTGTAAACAAGAGGGCC	149				
Db	25	CTTGTGGGCTCGGGCTCGACCTCGCTCGCGATGAAGTATAACAAGAGGGTC	84				
QY	150	AGGGGAGTGGTGGGGACAACATGGGCCCTGTGAGGCCTGTGGGTGCCCGCTTCCCCAG	209				
Db	85	GGGATGG-----CAGCGTAGGCCTGTGAGGCCTCGGGCTCGCGGTGCTGCCCCAG	134				
QY	210	CTCCCCCGCAGCCCGCTCCACAGTGTGCGCTCCGGTGTGGTGTACAGTGCCTATCGG	269				
Db	135	CTCCCCCGCAGCGGGCTCGCAGTGTGCACTCC-----GGTTGCCGGTGGGATTCGG	190				
QY	270	GTTCAGACCAAGGCTGCGTGTCTCCACCGCTTGTGT-----GGCCAGTGTT	319				
Db	191	GTTCGGACCAAGGCTGTGTCTCCGCGTTTATGTGGCCCCACAGGCGGGGTT	250				
QY	320	ACTGCGGTACGCCAGACAGAGCTCAGCCTATGGAGGAGCCTGTGTACCCCTCAGC	379				
Db	251	ACTGTGCGACCAAGACAGAGAGCTTTGGCGCTATGGAGGAGCCCGGGGCTACCCCTCAAC	310				
QY	380	CCTACCTGGGGCTGGTCTCGGAGAGCTAGCAGAGTTGTGGCAGCACTACCTGAGAGTA	439				
Db	311	CGGATTTGGGGCTGCTCCTCGAGAGAGCTACGACGGTTGTGGCAGCACTGCCTGAAGTA	370				
QY	440	TGACACAGATGAGAACTCCTTATGGTTTTCCATCGGAACCTGGTGGTATGTGCAGCTGTA	499				
Db	371	TGACACAGATCTAACTCTTATGGTTTTCCATGGGAAATGGTGATATGTGCAGCTGTTG	430				
QY	500	TTGATTTTGTGTTGTTCTCCTTTTTTGTGGAGAAGTTTATGATCGGTTAGGAGTCGGC	559				
Db	431	TTGATTTTGTGTTGTTCTCTTTTTTTGTGGAGAAGTTTATGATCGGTTAGGAGTCGGC	490				
QY	560	TTTACGTGGGAAGAGAGCAAAAATTGGTGCACGCTTCTCGGACTAATTGAAGAAAAAT	619				
Db	491	TTTATGTGGACGAGAGAAAAAGCTTGCTCTAATGCTTCTCGGACTAATTGAAGAAAAAA	550				
QY	620	GTAACCTACTTGAAAAATTTAGCCTTATTCAAAAGAGTAGTGAAGGCTATGAAGTAGAGT	679				
Db	551	GTAACCTACTTGAAAAATTTAGCCTTATTCAAAAGAGTAGTGAAGGCTATGAAGTAGAGT	610				
QY	680	CATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAA-----GAAGCAC	727				
Db	611	CATCTTTAAGGATGCCAGCTTTGAGAGGAGGAGCAACAGAGAGCACAAAGTTTGGAGGTAG	670				
QY	728	GAGTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAACTTCGAACTTGAGATGAAA	787				
Db	671	AAATATCAAAATCGCAACCTGTGAAAGCTGAACAGGTCCAACTTCGAACTTGAGATGAAA	730				
QY	788	TCCTCTGTCTAGAAAAAGACTTAAAAACAAGAGAAATCTAAAATTTCTCAACAAGATGAAT	847				

QY 1928 TGGGTTGGCCCTCATCTGAAACAGAGAGCTTTCTCTCTCTCCCTCCAACTTTGTTGGAGGGTC 1987
Db |||||
QY 1868 TGGGTTGGCCCTCATCTGAAACAGAGAGCTTTCTCTCTCTCCCTCCAACTTTGTTGGAGGGTC 1927
Db |||||
QY 1988 CACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGAGGCTCAGAGGCCACAGGA 2047
Db |||||
QY 1928 CACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGAGGCTCAGAGGCCACAGGA 1987
Db |||||
QY 2048 ATCCTCTGGACATCAGATTAACCAATGAAGAGGAGATCAAGCTGTGATAGGTTAACCG 2107
Db |||||
QY 1988 ATCCTCTGGACATCAGATTAACCAATGAAGAGGAGATCAAGCTGTGATAGGTTAACCG 2047
Db |||||
QY 2108 ATCCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGGGACCCAGACCGTA 2167
Db |||||
QY 2048 ATCCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGGGACCCAGACCGTA 2107
Db |||||
QY 2168 GGATGATGTTTCCCTCCGACGAGCAATCATATCTGATTCAGCCCTTCCCTCCAAAGGC 2227
Db |||||
QY 2108 GGATGATGTTTCCCTCCGACGAGCAATCATATCTGATTCAGCCCTTCCCTCCAAAGGC 2167
Db |||||
QY 2228 AAGACAGATTTTGTCTTAATCTGCTAGACTGTCTGGACCAAGAGAACTCAGAAAGTTTAA 2287
Db |||||
QY 2168 AAGACAGATTTTGTCTTAATCTGCTAGACTGTCTGGACCAAGAGAACTCAGAAAGTTTAA 2227
Db |||||
QY 2288 ATATCCCTTCTTTGGATTAATGAGTGGGTCAATGCTTCCAGAAATGGAATCCAGTAGAA 2347
Db |||||
QY 2228 ATATCCCTTCTTTGGATTAATGAGTGGGTCAATGCTTCCAGAAATGGAATCCAGTAGAA 2287
Db |||||
QY 2348 ATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCTGCTGAAA 2407
Db |||||
QY 2298 ATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCTGCTGAAA 2347
Db |||||
QY 2408 ATGAGGCACTGGCCCTGGCTTTGCTTCTCAGCTCTTGTCTCCAAATCAGAGGTCCATGT 2457
Db |||||
QY 2348 ATGAGGCACTGGCCCTGGCTTTGCTTCTCAGCTCTTGTCTCCAAATCAGAGGTCCATGT 2407
Db |||||
QY 2468 TTCCAGTGATGCAAGAGGCCCATCTTGGAGAGAGGACCTCTTCCCTCCCACTCCTC 2527
Db |||||
QY 2408 TTCCAGTGATGCAAGAGGCCCATCTTGGAGAGAGGACCTCTTCCCTCCCACTCCTC 2467
Db |||||
QY 2528 CAGGAGCCATGTTGGAGCTTCTGAGATTAATTTCCACCAAGGGATTTCCAGGTCCAC 2587
Db |||||
QY 2468 CAGGAGCCATGTTGGAGCTTCTGAGATTAATTTCCACCAAGGGATTTCCAGGTCCAC 2527
Db |||||
QY 2588 CACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTTACCTTC 2647
Db |||||
QY 2528 CACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTTACCTTC 2587
Db |||||
QY 2648 CCCCAGACCTGGATTTTCCCCCACCACCCACATTTCTGAAGGTAGAAGTGAGTTCCCT 2707
Db |||||
QY 2588 CCCCAGACCTGGATTTTCCCCCACCACCCACATTTCTGAAGGTAGAAGTGAGTTCCCT 2647
Db |||||
QY 2708 CAGGTTGATTCACCTTCAATGAGCTGCTACTGACATCCAGAACCCAGCAAGAA 2767
Db |||||
QY 2648 CAGGTTGATTCACCTTCAATGAGCTGCTACTGACATCCAGAACCCAGCAAGAA 2707
Db |||||
QY 2768 CCTGACAAATATTTTCTCTTCAAAAGTAATTTTACTGATCTCATTTTCAGTTTAAAG 2827
Db |||||
QY 2708 CCTGACAAATATTTTCTCTTCAAAAGTAATTTTACTGATCTCATTTTCAGTTTAAAG 2767
Db |||||
QY 2828 TAACTGCTGTTACTTAAGTATTAACCTTTTCTCAATTTGAAGCTTAATGGAATTATA 2887
Db |||||
QY 2768 TAACTGCTGTTACTTAAGTATTAACCTTTTCTCAATTTGAAGCTTAATGGAATTATA 2827
Db |||||
QY 2888 TTCTCAGGATAGTATTTTCTTAATTAAGATGATTTAAATATGAATCTTATGAGTAA 2944
Db |||||
QY 2828 TTCTCAGGATAGTATTTTCTTAATTAAGATGATTTAAATATGAATCTTATGAGTAA 2884
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:10:46 ; Search time 1155 Seconds
(without alignments)

11637.488 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttaatccatgaagaag.....acttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3164	100.0	3164	3 AAZ58979	Aaz58979 Human cyt
2	2860.4	91.0	3108	7 ABQ77437	Abq77437 Human CGD
3	2567.8	81.2	3681	5 AAS69467	Aas69467 DNA encod
4	2515.2	79.5	2928	7 ABQ77424	Abq77424 Human CGD
5	2501.4	79.1	3915	5 AAS84445	Aas84445 DNA encod
6	2483.4	78.5	3296	7 ACD19410	Ac119410 cDNA enco
7	2414.6	76.3	2814	6 ABL67999	Ab167999 Ovary can
8	2414.6	76.3	2814	6 ABL68239	Ab186239 Human PRO
9	2414.6	76.3	2814	6 ABL96922	Abn96922 Gene #342
10	2414.6	76.3	2814	6 ABL95728	Ab195728 Human ang
11	2414.6	76.3	2814	9 ADD10624	Add10624 Human sec
12	2414.6	76.3	2814	9 ADD11584	Add11584 Human sec
13	2414.6	76.3	2814	9 ADD37377	Add37377 Human sec
14	2414.6	76.3	2814	10 ADE41585	Ad41585 Human sec
15	2397.2	75.8	2602	7 ACD13199	Ac13199 cDNA enco
16	2384.6	75.4	2762	5 AAS69466	Aas69466 DNA encod
17	2326.2	73.5	4345	5 AAF61716	Aaf61716 Human CTA
18	2303.4	72.8	2614	7 ACD13200	Ac13200 cDNA enco
19	2292.2	72.4	2775	5 AAS84444	Aas84444 DNA encod
20	2159.8	68.3	2351	4 AAI24168	Aai24168 Probe #14
21	2159.8	68.3	2351	4 ABA69288	Aba69288 Human foe
22	2159.8	68.3	2351	4 AAI49455	Aai49455 Probe #18
23	2159.8	68.3	2351	4 ABA51280	Aba51280 Human bre

24	2159.8	68.3	2351	4 ABA36214	Aba36214 Probe #14
25	2159.8	68.3	2351	4 AAK17575	Aak17575 Human bra
26	2159.8	68.3	2351	4 ABS43008	Abs43008 Human liv
27	2159.8	68.3	2351	5 AAI09732	Aai09732 Probe #97
28	2159.8	68.3	2351	6 ABS17476	Ab17476 Human gen
29	2148.6	67.9	2334	4 AAI24172	Aai24172 Probe #14
30	2148.6	67.9	2334	4 ABA69293	Aba69293 Human foe
31	2148.6	67.9	2334	4 AAI49459	Aai49459 Human bre
32	2148.6	67.9	2334	4 ABA51285	Aba51285 Human bre
33	2148.6	67.9	2334	4 ABS43012	Ab143012 Human liv
34	2148.6	67.9	2334	6 ABA51285	Aba51285 Human bre
35	2148.2	67.9	3263	5 AAS84443	Aas84443 DNA encod
36	2111.4	66.7	2588	7 ABT16043	Abt16043 NOXV rela
37	2026.2	64.0	2483	7 ABT16044	Abt16044 NOXV rela
38	1954.6	61.8	4292	2 AAX80481	Aax80481 Human sec
39	1954.6	61.8	4292	2 AAS92227	Aas92227 Human CDN
40	1954.6	61.8	4292	6 ABA90896	Aba90896 Human pol
41	1838.8	58.1	1986	4 AAI4980	Aai4980 Probe #49
42	1838.8	58.1	1986	4 ABA56714	Aba56714 Human foe
43	1838.8	58.1	1986	4 AAI36326	Aai36326 Probe #50
44	1838.8	58.1	1986	4 ABA46168	Aba46168 Human bre
45	1838.8	58.1	1986	4 ABA26329	Aba26329 Probe #47

ALIGNMENTS

RESULT 1

AAZ58979

ID AAZ58979 standard; CDNA; 3164 BP.

XX AC AAZ58979;

DT 08-MAY-2000 (first entry)

DE Human cytoskeletal protein (HCYT) encoding cDNA (clone 3768043).

XX KW Human cytoskeletal protein; HCYT; cell proliferation; immunological;
KW reproductive; nervous disorder; cytostatic; immunosuppressive; anti-HIV;
KW anti-diabetic; antiarteriosclerotic; dermatological; anti-inflammatory;
KW anti-infertility; vasotropic; cardiant; ss.
XX Homo sapiens.
XX

XX FH Key Location/Qualifiers
XX CDS 352..2772
XX FT /*tag= a

XX PN WO200006730-A2.

XX PD 10-FEB-2000.

XX PF 30-JUL-1999; 99WO-US017167.

XX PR 31-JUL-1998; 98US-0155185P.

XX PR 04-AUG-1998; 98US-0160081P.

XX PR 19-AUG-1998; 98US-0155228P.

XX (INCY-) INCYTE PHARM INC.

XX PA Bandman O, Tang YT, Yue H, Corley NC, Guegler KJ, Azimzai Y;

XX PI Patterson C, Lal P, Baughn MR;

XX DR MPI; 2000-195297/17.

XX DR P-PSDB; AAY77574.

XX PT Human cytoskeletal proteins useful for diagnosing, treating preventing
XX cell proliferative, immunological, reproductive, developmental and
XX nervous disorders.
XX PS Claim 9; Page 79-80; 84pp; English.
XX CC The invention provides human cytoskeletal proteins (HCYT) and nucleic

CC acids encoding the proteins. The HCVYR polypeptides can be expressed by
 CC standard recombinant methodology. The HCVYR polypeptides, modulators and
 CC antibodies are useful for treating or preventing a disorder associated
 CC with decreased and increased expression or activity of HCVYR in mammals.
 CC The polypeptides are also useful for diagnosing HCVYR activity disorders
 CC such as cell proliferative, immunological, reproductive, developmental
 CC and nervous disorders. Sequences AA258974-991 represent cDNA clones
 CC encoding the HCVYR polypeptides
 XX

SQ Sequence 3164 BP; 1015 A; 632 C; 685 G; 832 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3164; DB 3; Length 3164;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAGCTT 60
 DB 1 AGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAGCTT 60
 QY 61 CAAATATGTGATGAAACTGCCAGAACTTAAGCGCGCGGCTCAGACAGCGTGC 120
 DB 61 CAAATATGTGATGAAACTGCCAGAACTTAAGCGCGCGGCTCAGACAGCGTGC 120
 QY 121 CTCAGGATGTAAGTGTAAAGAGAGCGCAGGGAGGTGTGGGGGACAAATCGGCGCTG 180
 DB 121 CTCAGGATGTAAGTGTAAAGAGAGCGCAGGGAGGTGTGGGGGACAAATCGGCGCTG 180
 QY 181 TGAGCCGTGTGGTCCCGGCTTCCAGCTCCCGCCAGCGCTCCACAGTGTCCG 240
 DB 181 TGAGCCGTGTGGTCCCGGCTTCCAGCTCCCGCCAGCGCTCCACAGTGTCCG 240
 QY 241 CTCGGTGTGGTGTCAAGTGTGGGCTTCCAGACCCCAAGGCTGGTGTCTCCACC 300
 DB 241 CTCGGTGTGGTGTCAAGTGTGGGCTTCCAGACCCCAAGGCTGGTGTCTCCACC 300
 QY 301 GCTTGTGTGGCAGTGTACTCGCGTGAAGCGCAGCAGCGCTCGAGCTATGGAGAG 360
 DB 301 GCTTGTGTGGCAGTGTACTCGCGTGAAGCGCAGCAGCGCTCGAGCTATGGAGAG 360
 QY 361 CTGTGTGTACCCCTCAGCCCTACCTGGGCTGTGGTGTGGAGGAGTACGAGAGTGTG 420
 DB 361 CTGTGTGTACCCCTCAGCCCTACCTGGGCTGTGGTGTGGAGGAGTACGAGAGTGTG 420
 QY 421 GCAGCACTACCTGAGAGTATGACACAGATGAGAACTTTATGTTTCCATCGAACTG 480
 DB 421 GCAGCACTACCTGAGAGTATGACACAGATGAGAACTTTATGTTTCCATCGAACTG 480
 QY 481 GTGGTATGTGAGCTGTATTGATTTTGTGTTCTCCCTTTTGTGGAGAGTGT 540
 DB 481 GTGGTATGTGAGCTGTATTGATTTTGTGTTCTCCCTTTTGTGGAGAGTGT 540
 QY 541 AGATCGGTAGAGTCGGCTTTACGTGGAGAGAGCAAAACTTTGGTGCACCGCTTCT 600
 DB 541 AGATCGGTAGAGTCGGCTTTACGTGGAGAGAGCAAAACTTTGGTGCACCGCTTCT 600
 QY 601 GGACTAATGAGAAATGAACTTAACTTGAATAATTTAGCTTATTCAAAAGAGTAT 660
 DB 601 GGACTAATGAGAAATGAACTTAACTTGAATAATTTAGCTTATTCAAAAGAGTAT 660
 QY 661 GAAGCTATGAGTATGAGTATCTTTAGAGGATGCGAGCTTTGAGAAGGAGGAGAGAA 720
 DB 661 GAAGCTATGAGTATGAGTATCTTTAGAGGATGCGAGCTTTGAGAAGGAGGAGAGAA 720
 QY 721 GAAGCAGAGGTTTGGAGGACCTGTGAAGCTGACAGCTTCAATCTGAACTTGG 780
 DB 721 GAAGCAGAGGTTTGGAGGACCTGTGAAGCTGACAGCTTCAATCTGAACTTGG 780
 QY 781 GATGAATCTCTGCTAGAAAAAGACTTAAACCAAGAGAAATCTAAACATTTCTCAACA 840
 DB 781 GATGAATCTCTGCTAGAAAAAGACTTAAACCAAGAGAAATCTAAACATTTCTCAACA 840
 QY 841 GATGAATGATGGCGGATATTTCAAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
 DB 841 GATGAATGATGGCGGATATTTCAAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900

DB 841 GATGAATGATGGCGGATATTTCAAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
 QY 901 CTCAAATCAAAATAGCTGAAGCCCAAAATCATCTCAAGACATTTTAAATAGTCAAGAA 960
 DB 901 CTCAAATCAAAATAGCTGAAGCCCAAAATCATCTCAAGACATTTTAAATAGTCAAGAA 960
 QY 961 CGACGGCTTATAGCAATAAAGATGCTTTGAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1020
 DB 961 CGACGGCTTATAGCAATAAAGATGCTTTGAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1020
 QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
 DB 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
 QY 1081 CAGAAAATAAATTTGAAGACTCCAAAGTACACGCGAGAACAAAGTCTGAATGATAAGAA 1140
 DB 1081 CAGAAAATAAATTTGAAGACTCCAAAGTACACGCGAGAACAAAGTCTGAATGATAAGAA 1140
 QY 1141 AATCATCATCAAGACCTGACTCGACATTTGCCAATGATGAAGATCAGGCTCTGTGCTT 1200
 DB 1141 AATCATCATCAAGACCTGACTCGACATTTGCCAATGATGAAGATCAGGCTCTGTGCTT 1200
 QY 1201 GAAGAGACACAAACGAGTATTAACCTGGAAATTTAGAAAGTGAACAGTGAATCGGAAAT 1260
 DB 1201 GAAGAGACACAAACGAGTATTAACCTGGAAATTTAGAAAGTGAACAGTGAATCGGAAAT 1260
 QY 1261 GGTGCTTACTAGATAATCTCCAAAGAGGAGCTTTGAAGAACTGATTCATGCTCTGAAG 1320
 DB 1261 GGTGCTTACTAGATAATCTCCAAAGAGGAGCTTTGAAGAACTGATTCATGCTCTGAAG 1320
 QY 1321 TTAATATCTCTTTTAAAAACCTTTAGAAAGAGAGAAACCAAAATTTATTTCAAGTGTCT 1380
 DB 1321 TTAATATCTCTTTTAAAAACCTTTAGAAAGAGAGAAACCAAAATTTATTTCAAGTGTCT 1380
 QY 1381 GAAGTGTATTAACAAAGAGAGCTTACAGAGATATTAATAATCTTCAGACTCAACAA 1440
 DB 1381 GAAGTGTATTAACAAAGAGAGCTTACAGAGATATTAATAATCTTCAGACTCAACAA 1440
 QY 1441 GCATCTTTCCAGTCAGAAACACACATTTTGAATAATGAGAACTCAGAACTTCAACAGAAA 1500
 DB 1441 GCATCTTTCCAGTCAGAAACACACATTTTGAATAATGAGAACTCAGAACTTCAACAGAAA 1500
 QY 1501 CTTAAAGTATGACTGAATTTATCAAGAAATGAAATGAAACTCCACAGGAAATTAACA 1560
 DB 1501 CTTAAAGTATGACTGAATTTATCAAGAAATGAAATGAAACTCCACAGGAAATTAACA 1560
 QY 1561 GTAGAGAAAATTTATCGGTTAGAGAAAGAGAAACCTTTCTAAAGTAGATGAAGATC 1620
 DB 1561 GTAGAGAAAATTTATCGGTTAGAGAAAGAGAAACCTTTCTAAAGTAGATGAAGATC 1620
 QY 1621 AGCCATGCCACTGAAGAGCTGAGACCTATAGAAAAGGAGGCAAAAGATCTTGAAGAGAA 1680
 DB 1621 AGCCATGCCACTGAAGAGCTGAGACCTATAGAAAAGGAGGCAAAAGATCTTGAAGAGAA 1680
 QY 1681 TTGAGAGAACTTATCTTATTTCAAGGCGAGATTTTCCATGAGAAAAAGACAT 1740
 DB 1681 TTGAGAGAACTTATCTTATTTCAAGGCGAGATTTTCCATGAGAAAAAGACAT 1740
 QY 1741 GATAATTGGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
 DB 1741 GATAATTGGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
 QY 1801 GCTCAACACAGACAAAATTTAATCTGAACAGAGCTTAAATTTGAACTTTAGAAAAAGAT 1860
 DB 1801 GCTCAACACAGACAAAATTTAATCTGAACAGAGCTTAAATTTGAACTTTAGAAAAAGAT 1860
 QY 1861 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCAATCCCATATGCTCC 1920
 DB 1861 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCAATCCCATATGCTCC 1920
 QY 1921 TCACCAATGGGTTGGCCTTCACTGAAACAGAGCTTTTCTCTCTCCCTCCACTTTGTTG 1980
 DB 1921 TCACCAATGGGTTGGCCTTCACTGAAACAGAGCTTTTCTCTCTCCCTCCACTTTGTTG 1980

1981 GAGGTCCTCAGACTCTCCTCTTCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
Db |
1981 GAGGTCCTCAGACTCTCCTCTTCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
Qy |
2041 CCAGGAATCCTCTGACCATCAGATTACCAATGAAGAGGAGAAATCAAGCTGTATAGG 2100
Db |
2041 CCAGGAATCCTCTGACCATCAGATTACCAATGAAGAGGAGAAATCAAGCTGTATAGG 2100
Qy |
2101 TTAAACGATCCTCATAGGCTCCTCTGACACTGGGTCTCTGTCACCTCCATGGGACGAG 2160
Db |
2101 TTAAACGATCCTCATAGGCTCCTCTGACACTGGGTCTCTGTCACCTCCATGGGACGAG 2160
Qy |
2161 GACCGTAGGATGATGTTCTCCGCGCAGGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
Db |
2161 GACCGTAGGATGATGTTCTCCGCGCAGGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
Qy |
2221 CAAGGCAAGACAGATTTGTTCTTAATCTGAGACTGGGTCTCTGTCACCTCCATGGGACGAG 2280
Db |
2221 CAAGGCAAGACAGATTTGTTCTTAATCTGAGACTGGGTCTCTGTCACCTCCATGGGACGAG 2280
Qy |
2281 AGTTTAAATATGCTTCTTGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAATCC 2340
Db |
2281 AGTTTAAATATGCTTCTTGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAATCC 2340
Qy |
2341 AGTAGAAATGATACCAAGATGATCTTGGTAATTTAAATGTGCTGATTCATCTCTCCT 2400
Db |
2341 AGTAGAAATGATACCAAGATGATCTTGGTAATTTAAATGTGCTGATTCATCTCTCCT 2400
Qy |
2401 GCTGAAATGAAGCCACTGGCCCTGCTTGTCTTCCCTCCACTCTTCTCTCAATCAGAGGT 2460
Db |
2401 GCTGAAATGAAGCCACTGGCCCTGCTTGTCTTCCCTCCACTCTTCTCTCAATCAGAGGT 2460
Qy |
2461 CCATGTTTCCAGTGATGACAGAGGCCATTTCTGAGAGGAGGACTCTCTTCCCTCCCA 2520
Db |
2461 CCATGTTTCCAGTGATGACAGAGGCCATTTCTGAGAGGAGGACTCTCTTCCCTCCCA 2520
Qy |
2521 CTTCTCCAGAGGCCATTTGAGGCTTCTGAGATTTTTCACCAAGAGGATTTCCCA 2580
Db |
2521 CTTCTCCAGAGGCCATTTGAGGCTTCTGAGATTTTTCACCAAGAGGATTTCCCA 2580
Qy |
2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTTCTCTCT 2640
Db |
2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTTCTCTCT 2640
Qy |
2641 TACCTTCCCAAGACCTGATTTTCCCTCCACCCACCCATTTCTGAGTGAAGTGAAG 2700
Db |
2641 TACCTTCCCAAGACCTGATTTTCCCTCCACCCACCCATTTCTGAGTGAAGTGAAG 2700
Qy |
2701 TTCCCTCAGGTTTGAATCCACCTTCAATGAGCCTGCTACTGAACATCCAGAACACAG 2760
Db |
2701 TTCCCTCAGGTTTGAATCCACCTTCAATGAGCCTGCTACTGAACATCCAGAACACAG 2760
Qy |
2761 CAAGAACCTGACATATTTTGTCTCTTCAAAAGTAATTTGACTGATCTCATTTCA 2820
Db |
2761 CAAGAACCTGACATATTTTGTCTCTTCAAAAGTAATTTGACTGATCTCATTTCA 2820
Qy |
2821 GTTTAAGTAAGTCTGTTACTTAAGTATGATTAACCTTTGCTCAAAATGAAGCTTAATGA 2880
Db |
2821 GTTTAAGTAAGTCTGTTACTTAAGTATGATTAACCTTTGCTCAAAATGAAGCTTAATGA 2880
Qy |
2881 ATTATAATCTCAGATAGTATTTTGAATAAAGATGATTTAAATGATGATCTTATGAG 2940
Db |
2881 ATTATAATCTCAGATAGTATTTTGAATAAAGATGATTTAAATGATGATCTTATGAG 2940
Qy |
2941 TAAATTTTCAATTTTATTTTACAGGATTAATCAATTTCAATTTGATTTAATCCACTATT 3000
Db |
2941 TAAATTTTCAATTTTATTTTACAGGATTAATCAATTTCAATTTGATTTAATCCACTATT 3000
Qy |
3001 ATATAAACAATAGTGGAGTTTATATATGATTAATGTAATCTTTCAGGTGGGAGGCTTTAATTC 3060
Db |
3001 ATATAAACAATAGTGGAGTTTATATATGATTAATCTTTCAGGTGGGAGGCTTTAATTC 3060

3061 TGAAGTCTGTCTTTATGTCACCAAGAACTGTATTTACTGTGGTTGTGCAAAATGTCAAAG 3120
3061 TGAAGTCTGTCTTTATGTCACCAAGAACTGTATTTACTGTGGTTGTGCAAAATGTCAAAG 3120
3121 TAACCTTATGCTTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3164
3121 TAACCTTATGCTTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3164

RESULT 2
ABQ77437
ID ABQ77437 standard; cDNA; 3108 BP.
XX
AC ABQ77437;
DT 10-MAY-2003 (first entry)
XX Human CGDD cDNA 7506167CB1 SEQ ID 51.
XX Human; cell growth, differentiation and death; CGDD; cardiact; cancer;
KW cytosolic; neuroprotective; gene therapy; diagnosis; cardiovascular;
KW neurological disorder; gene; ss.
OS Homo sapiens.
XX
PN WO2003014322-A2.
XX
PD 20-FEB-2003.
XX
PF 08-AUG-2002; 2002WO-US025465.
XX
PR 08-AUG-2001; 2001US-0311017P.
PR 17-AUG-2001; 2001US-0313070P.
PR 17-AUG-2001; 2001US-0313071P.
PR 24-AUG-2001; 2001US-0314678P.
PR 31-AUG-2001; 2001US-0316692P.
PR 07-SEP-2001; 2001US-0317913P.
PR 14-SEP-2001; 2001US-0322182P.
PR 07-DEC-2001; 2001US-0340747P.
PR 20-DEC-2001; 2001US-0342761P.
PR 29-MAR-2002; 2002US-0369129P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;
PI Dugan BM, Elliott VS, Forsythe J, Gietzen KJ;
PI Gervad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
PI Kable AB, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
PI Sprague WW, Swarnakar A, Tang Y, Warren BA, Xu Y, Yao MG, Yue H;
PI Yue H;
XX
XX WPI; 2003-268197/26.
DR P-PSDB; ABG74698.
XX
XX New polypeptide associated with cell growth, differentiation and death,
PT useful for preparing a composition for diagnosing or treating a disease
PT e.g., cardiovascular or neurological disorders or cancer.
XX
XX Claim 106; Page 264-265; 267pp; English.
XX
XX This invention describes a novel polypeptide associated with cell growth,
CC differentiation and death (CGDD) which has cardiact, cytosolic and
CC neuroprotective activity. The polypeptides of the invention are useful
CC for preparing a composition for diagnosing, or treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional CGDD e.g., cardiovascular or neurological disorders or cancer.
CC The polynucleotides of the invention can be used for gene therapy.
CC ABQ77414-ABQ77441 encode the human CGDD polypeptides represented in
CC ABG74675-ABG74701
XX
SQ Sequence 3108 BP; 1011 A; 607 C; 673 G; 817 T; 0 U; 0 Other;
91.0%; Score 2880.4; DS 7; Length 3108;

Query Match

Best Local Similarity 95.7%; Pred. No. 0;			
Matches 3023; Conservative 0; Mismatches 6; Indels 129; Gaps 1;			
QY	1	AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAGCTT	60
Db	72	AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAGCTT	131
QY	61	CAAAATATGTGATGTGAAGAACTGCCAGAACTGAAGCGCGCGCGGTTCAGACAGCGCTGC	120
Db	132	CAAAATATGTGATGTGAAGAACTGCCAGAACTGAAGCGCGCGCGGTTCAGACAGCGCTGC	191
QY	121	CTCAGATGTAAAGTGTAAAGAGGGCCAGGGAGGTGTGGGGGCAAAATGGGCGCTG	180
Db	192	CTCAGATGTAAAGTGTAAAGAGGGCCAGGGAGGTGTGGGGGCAAAATGGGCGCTG	251
QY	181	TCAGGCGCTGTGGGTGCCCGCGTTCGCCAGTCCCGCGCGCGCTCCACAGTGGTCCG	240
Db	252	TCAGGCGCTGTGGGTGCCCGCGTTCGCCAGTCCCGCGCGCGCTCCACAGTGGTCCG	311
QY	241	CTCCGGTGGTGTCAAGTGGCGATTCGGGTTCAGAGCCCAAGGCTGGGTTCCTCCACC	300
Db	312	CTCCGGTGGTGTCAAGTGGCGATTCGGGTTCAGAGCCCAAGGCTGGGTTCCTCCACC	371
QY	301	GCTGTGTGGCGAGTGTACTGGGTGACCGCCAGAGCAGCTTCGAGCTATGAGGAG	360
Db	372	GCTGTGTGGCGAGTGTACTGGGTGACCGCCAGAGCAGCTTCGAGCTATGAGGAG	431
QY	361	CTGFGCTACCCCTCAGCCCTACCTCGGGCTGGTCTGAGGAGCTACGAGAGTTGTG	420
Db	432	CTGFGCTACCCCTCAGCCCTACCTCGGGCTGGTCTGAGGAGCTACGAGAGTTGTG	491
QY	421	GCAGCACTACCTGAGAGTATGACAGATGAGATCCTTTATGGTTTTCATCGGAACCTG	480
Db	492	GCAGCACTACCTGAGAGTATGACAGATGAGATCCTTTATGGTTTTCATCGGAACCTG	551
QY	481	GTGGTATGTCAGCTGTATGGATTTTGTGTGTCTCCTTTTGTGGAGAGTTT	540
Db	552	GTGGTATGTCAGCTGTATGGATTTTGTGTGTCTCCTTTTGTGGAGAGTTT	611
QY	541	AGATCGGTAGGAGTCGGCTTTACGTGGAGAGAGCAAACTTGGTCCACGCTTCT	600
Db	612	AGATCGGTAGGAGTCGGCTTTACGTGGAGAGAGCAAACTTGGTCCACGCTTCT	671
QY	601	GGACTAATGAAGAAAAATGTAACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT	660
Db	672	GGACTAATGAAGAAAAATGTAACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT	731
QY	661	GAAGGCTATGAAGTATGATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGNA	720
Db	732	GAAGGCTATGAAGTATGATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGNA	791
QY	721	GAAGCAGAAATTTGGAGGCAACCTGTGAAAAGCTGAAAGGTCCCAATTTCTGAACTTGAG	780
Db	792	GAAGCAGAAATTTGGAGGCAACCTGTGAAAAGCTGAAAGGTCCCAATTTCTGAACTTGAG	851
QY	781	GATGAAATCTCTGTCTAGAAAAAGACTTAAACAGAGAAATCTTAAACATTTCTCAACAA	840
Db	852	GATGAAATCTCTGTCTAGAAAAAGACTTAAACAGAGAAATCTTAAACATTTCTCAACAA	911
QY	841	GATGAAATTTGATGGCGGATATTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC	900
Db	912	GATGAAATTTGATGGCGGATATTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC	971
QY	901	CTCAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA	960
Db	972	CTCAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA	1031
QY	961	CGAGGGCTATGCAATTAAGAGTGTCTTCAATGAAATTTCTCACTTCAGACCAAGCCAT	1020
Db	1032	CGAGGGCTATGCAATTAAGAGTGTCTTCAATGAAATTTCTCACTTCAGACCAAGCCAT	1091
QY	1021	AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGAGTGAAGTGAATTAATAA	1080

Db	1092	AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGAGTGAAGTGAATTAATAA	1151
QY	1081	CAGAAATTAACATTTTGAAGACTCCAAAGTACACGAGAACCAAGTTCTGATGATAAGAA	1140
Db	1152	CAGAAATTAACATTTTGAAGACTCCAAAGTACACGAGAACCAAGTTCTGATGATAAGAA	1211
QY	1141	AATCACATCAAGACCTGTGACACTTGGCAACTTGGCAATGATGAAAGATCAGGCTGTGCTT	1200
Db	1212	AATCACATCAAGACCTGTGACACTTGGCAACTTGGCAATGATGAAAGATCAGGCTGTGCTT	1271
QY	1201	GAAGAGACACACGAGTATGATTAACCTGGAAATAGAAAGTGAACAGTCAATCGGAAAT	1260
Db	1272	GAAGAGACACACGAGTATGATTAACCTGGAAATAGAAAGTGAACAGTCAATCGGAAAT	1331
QY	1261	GGTGTCTTACTTAGATAATCTCCAAAGAGAGCTTTCAAGAAACTGATTCATGCTGCTAAG	1320
Db	1332	GGTGTCTTACTTAGATAATCTCCAAAGAGAGCTTTCAAGAAACTGATTCATGCTGCTAAG	1391
QY	1321	TTAAATGCTCTTTTAAACCTTTAGAGGAGAGAGAAACCAATTTATATTCAGTTGTCT	1380
Db	1392	TTAAATGCTCTTTTAAACCTTTAGAGGAGAGAGAAACCAATTTATATTCAGTTGTCT	1451
QY	1381	GAAGTTGATAAAAACAAAGGAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	1440
Db	1452	GAAGTTGATAAAAACAAAGGAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	1511
QY	1441	GCATCTTTCAGTTCAGAAACACACATTTTGAAGTGAAGATCAGAGCTTCAACAGAA	1500
Db	1512	GCATCTTTCAGTTCAGAAACACACATTTTGAAGTGAAGATCAGAGCTTCAACAGAA	1571
QY	1501	CTTAAAGTAAATGATCTGAATTTATCAAGAAAAATGAAATGAAATCTCCACAGAAATTAACA	1560
Db	1572	CTTAAAGTAAATGATCTGAATTTATCAAGAAAAATGAAATGAAATCTCCACAGAAATTAACA	1631
QY	1561	GTAGAGAAATTTATCGTTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGATC	1620
Db	1632	GTAGAGAAATTTATCGTTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGATC	1691
QY	1621	AGCCATCCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCCAAGATCTTGAAGAGAA	1680
Db	1692	AGCCATCCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCCAAGATCTTGAAGAGAA	1751
QY	1681	TTGAGAGAACTATTCATCTTATCAAGGCGAGATTTTCCCATGAGAAAAAGCAGAT	1740
Db	1752	TTGAGAGAACTATTCATCTTATCAAGGCGAGATTTTCCCATGAGAAAAAGCAGAT	1811
QY	1741	GATAATTTGGTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAAGAAAGAAAAAT	1800
Db	1812	GATAATTTGGTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAAGAAAGAAAAAT	1871
QY	1801	GCTCAACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1860
Db	1872	GCTCAACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1931
QY	1861	CCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCCTCATATGCTCC	1920
Db	1932	CCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCCTCATATGCTCC	1967
QY	1921	TCACCATTTGGTGGCCCTTCATCTGAAAACAAAGAGCTTTCTCTCTCTCTCAATTTGTTG	1980
Db	1968	TCACCATTTGGTGGCCCTTCATCTGAAAACAAAGAGCTTTCTCTCTCTCTCAATTTGTTG	1967
QY	1981	GAGGTCCTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGCTCACGAGGC	2040
Db	1968	GAGGTCCTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGCTCACGAGGC	1982
QY	2041	CCAGGGAATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAGG	2100
Db	1983	CCAGGGAATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAGG	2042
QY	2101	TTAACCGATCTCTAGGCTCCCTCTGACACTGGGTCTCTGTACCTCCCTCCAGGACAG	2160
Db	2043	TTAACCGATCTCTAGGCTCCCTCTGACACTGGGTCTCTGTACCTCCCTCCAGGACAG	2102

XX	AA569467;
AC	13-FEB-2002 (first entry)
XX	DNA encoding novel human diagnostic protein #5271.
DT	Human; chromosome mapping; gene mapping; gene therapy; forensic;
DE	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	Homo sapiens.
OS	WO200175067-A2.
XX	11-OCT-2001.
PN	30-MAR-2001; 2001WO-US008631.
XX	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
PP	(HYSE-) HYSEQ INC.
XX	Drmanac RT, Liu C, Tang YT;
PI	WPI: 2001-639362/73.
DR	P-PSDB; ABG05280.
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
PS	Claim 1; SEQ ID NO 5271; 103pp; English.
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 3681 BP; 1134 A; 685 C; 805 G; 1056 T; 0 U; 1 Other;
	Query Match 81.2%; Score 2567.8; DB 5; Length 3681;
	Best Local Similarity 92.9%; Pred. No. 0;
	Matches 2881; Conservative 0; Mismatches 168; Indels 53; Gaps 16
Qy	90 CTAAGCGGGCGGGGTCTCAGACACGCCTCCCTCAGATGTAAAGTGTAAACAAGAGGCC 149
Db	57 CTTGTTCGGGTTCGGGTTCGGACCTCCCTCCCTCGGGATGTAAAGTATTAAACAAGAGGTTC 116
Qy	150 AGCGGAGGTGGTGGGGACCAATCGGCCTGTGAGGCCTGTGGGTGCCCGCGTTCCTCCAG 209
Db	117 GGAGTGG-----CACCGTAGGCTGTGAGGCTTCGGGTGCCCTGTCTCCCCAG 166
Qy	210 CTCCTCCCAGCCCGCTCCAACGTGGTCCGCTCCGTTGGTTGTACGTGGCATTCGG 269

QY 210 CTCCCCCGCAGCCCGCTCCACAGTGGTCCGCTCCGGTGGTGTGTCACGTGGCATTCG 269

Db 167 CTCCCGGCGAGCCGGCTCCCGAGTGGTCCACTCC-----GGTTCGGGTGCGGATTCGG 222
Qy 270 GTTCCAGACCCAGAGCTGGTGTCTCCACCGCTGTGTGT-----GGCCAGTGT 319
Db 223 GTTCCGAGCCGAGAGCTGTGTGTCTCCGCGGTCAATTTGTGGCCCCGACAGCCGGGTT 282
Qy 320 ACTGGGTGACCGCCGAGAGCAGCTTCGACGCTATCGAGAGCGCTGGTGTCTACCCCTCAGC 379
Db 283 ACTGTGGCGACACAGAGAGCAGCTTTGGCGCTATGGAGGAGCCCGGGGCTACCCCTCAAC 342
Qy 380 CTTACTGGGGGTGGTCTGGAGGAGCTACGAGAGTTGTGCGAGCACTACTCGAGAGTA 439
Db 343 CGTATTTGGGGGTGCTCTCGAGAGAGCTACGAGCGGTGTGGCAGCACTGCTGAGGTA 402
Qy 440 TGAGACAGATGAGAAATCCTTATGTTTCCATCGGAACCTGGTGTATGTGCGAGCTGTTA 499
Db 403 TGAGACAGATGCTAATCTTTATGTTTTCATGGAAGTTTGTGATATGTGCGAGCTGTTG 462
Qy 500 TTTGGATTTTGTGTCTCTCTTTTGTGGAGAGTTTGTAGTCGGTTAGGAGTCGCG 559
Db 463 TTGGATTTTGTGTCTCTCTTTTGTGGAGAGTTTGTAGTCGGTTAGGAGTCGCG 522
Qy 560 TTTTACGTGGAGAGAGCAAAACCTTGTGTCAACGCTTTCTGACTAATTTGAAGAAAAAT 619
Db 523 TTTATGTGGAGAGAGAGAAAGCTTGTCTTAATGCTTTCTGACTAATTTGAAGAAAAA 582
Qy 620 GTAACACTCTGAAATTTAGCCTTATTCAAAAGAGTATGAGGCTATGAGTAGAGT 679
Db 583 GTAACACTCTGAAATTTAGCCTTATTCAAAAGAGTATGAGGCTATGAGTAGAGT 642
Qy 680 CATCTTTAGAGTAGCGAGCTTTGAGAGGAGGCGAGAGAGACAGAACTTTGGAGG 739
Db 643 CATCTTTAAGAGTAGCGAGCTTTGAGAGGAGGCGAAC---AGAAGCACAAAGTTTGGAGG 699
Qy 740 CAACCTGTGAAGCTGAACAGCTCAATCTGAACTTGAGGATGAAATCTCTCTCTAG 799
Db 700 CAACCTGTGAAGAGCTGAACAGCTCAATCTGAACTTGAGGATGAAATCTCTCTCTAG 759
Qy 800 AAAAAAGCTTTAAACAGAGAAATCTAAACATTTCTCAACAGATGAATTTGATGGCGGATA 859
Db 760 AAAAAAGCTTTAAACAGAGAAATCTAAACATTTCTCAACAGATGAATTTGATGGCGGATA 819
Qy 860 TTTCAAAAGTATACAGCTCTAGAGATGAGTCAAAATCCCTCAAAATCACAATAGCTG 919
Db 820 TTTCAAAAGGATACAGCTCTAGAGATGAGTCAAAATCCCTCAAAATCACAATAGCTG 879
Qy 920 AAGCCAAATCATCTGCAACACATTTAAATAGTGAAGAACGACGGGCTATAGCAATAA 979
Db 880 AAGCCAAATGACCTTCCAGATATTTCCATGAATGAAGAACGACCTGAGATAGCAATAA 939
Qy 980 AAGATGCTTTGAATGAATAATCTCACTTCAGACAGCCATTAACAGCTTTTTCAGCAAG 1039
Db 940 AAGATGCTTTGAATGAATAATCTCACTTCAGGAAAGCCAGAAACAGCTTTT---GCAAG 996
Qy 1040 AAGCTGAAGTATGAAAGGAGAGTGAAGTCAACTTAATAACAGAAATAACATTTGAAG 1099
Db 997 AAGCTGAAGTATGAAAGAGAACAGTGAAGTGAATTAATAACAGAAATAACATTTGAAG 1056
Qy 1100 ACTCCAAAGTACCGCAGAAACAAGTCTGAATGATAAAGAAATCAATCAAGACCCCTGA 1159
Db 1057 ACTCCAAAGTACATGCAAGAACAAAGTCTTAATGATAAAGAAAGTCAATCAAGACTCTGA 1116
Qy 1160 CTGACACTTTGCCATGATGAAGATCAGGCTGCTGTGTTGAAGAGACACACCGGATG 1219
Db 1117 CTGAACGCTTTTGAAGATGAAGATTTGGGCTGCTATGCTTTGGAGAGACATTAACCGATG 1176
Qy 1220 ATGATAACCTGGAAATTAAGAGTGAACAGTGAATCGGAAATTTGCTTACTTTAGATAATC 1279
Db 1177 ATGATAACTTTGGAAATTAAGAAATGAACAGTGAATCGGAAATTTGCTTACTTTAGATAATC 1236
Qy 1280 CTCACAAAGAGCTTTGAAGAACTGATTCATGCTGCTAGTTAAATGCTTTTAAAAA 1339
Db 1237 CTCACAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAGTTAAATGCTTTTAAAAA 1296

Qy 1340 CTTTAGAAGGAGAGAAACCAAAATTTATTCAGTTGTCTGAAGTTGATATAAAACAAAGG 1399
Db 1297 CTTTAGAAGGAGAGAAACCAAAATTTATTCAGTTGTCTGAAGTTGATATAAAACAAAGG 1356
Qy 1400 AAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAGCATCTTTGCAAGTCAAGAA 1459
Db 1357 AAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAGCATCTTTGCAAGTCAAGAA 1416
Qy 1460 ACACACATTTTGAATAATGAGAACTCAGAGCTTCAACAGAAACTTAAAGTAATGACTGAAT 1519
Db 1417 ACACACATTTTGAATAATGAGAACTCAGAGCTTCAACAGAAACTTAAAGTAATGACTGAAT 1476
Qy 1520 TATATCAAGAAATGAATGAATGAATCCACAGAAATTTAACAGTAGAGGAAATTTATTCGGT 1579
Db 1477 TATATCAAGAAATGAATGAATGAATCCACAGAAATTTAACAGTAGAGGAAATTTATTCGGT 1536
Qy 1580 TAGAGAAAGAGAGAAACTTTCTAAAGTAGATGAATAAGATCAGCCATGCGCATGAAGAGC 1639
Db 1537 TAGAGAAAGAGAGAAACTTTCTAAAGTAGATGAATAAGATCAGCCATGCGCATGAAGAGC 1596
Qy 1640 TGGAGACCTATAGAAAGCGGACCAAGATCTTGAAGAAGATTTGGAGAGAACTTATTCAT 1699
Db 1597 TGGAGACCTATAGAAAGCGGACCAAGATCTTGAAGAAGATTTGGAGAGAACTTATTCAT 1656
Qy 1700 CTTATCAAGGGCAGATTTATTTCCATGAGAAAAAGCACATGATAATTTGGTGGCAGCTC 1759
Db 1657 CTTATCAAGGGCAGATTTATTTCCATGAGAAAAAGCACATGATAATTTGGTGGCAGCTC 1716
Qy 1760 GGAATGCTGAAAGAACTCAATGATTTAAGNAAGAAATGCTCACAACAGACAAAAAT 1819
Db 1717 GGAATGCTGAAAGAACTCAATGATTTAAGNAAGAAATGCTCACAACAGACAAAAAT 1776
Qy 1820 TAACCTGAAACAGAGCTTTAAATTTGAACTTTTAGAAAAAGATCCTTATGCACTCGATGTTT 1879
Db 1777 TAACCTGAAACAGAGCTTTAAATTTGAACTTTTAGAAAAAGATCCTTATGCACTCGATGTTT 1836
Qy 1880 CAATAACAGATTTGGCAGAGAGATTTCCCATATGGTCCCTCACCATTGGGTTGGCCTT 1939
Db 1837 CAATAACAGATTTGGCAGAGAGATTTCCCATATGGTCCCTCACCATTGGGTTGGCCTT 1896
Qy 1940 CATCTGAAACAGAGCTTTTCTCTCTCTCCAACTTTGTTGGAGGTTCCACTCAGACTCT 1999
Db 1897 CATCTGAAACAGAGCTTTTCTCTCTCTCCAACTTTGTTGGAGGTTCCACTCAGACTCT 1956
Qy 2000 CACCTTTGCTTTCAGGGGAGGAGAGAGCTCAAGAGCTGTGATAGTTAAACCGATCCTCATAGG 2119
Db 1957 CACCTTTGCTTTCAGGGGAGGAGAGAGCTCAAGAGCTGTGATAGTTAAACCGATCCTCATAGG 2076
Qy 2060 ATCAGATTTACCAATGAAGAGAGAGATCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2179
Db 2017 ATCAGATTTACCAATGAAGAGAGAGATCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2076
Qy 2120 CTCCTCTGACACTGGGTCTCTGTCCTCATGGGACAGGACCGTAGGATGATGTTTTC 2179
Db 2077 CTCCTCTGACACTGGGTCTCTGTCCTCATGGGACAGGACCGTAGGATGATGTTTTC 2136
Qy 2180 CTCGGCAGAGCAATCATATCTGATTCAGCCTTCTCTCCACAAAGCAAGACAGATTTT 2239
Db 2137 CTCGGCAGAGCAATCATATCTGATTTGAGCCTTCTCTCCACAAAGCAAGACAGATTTT 2196
Qy 2240 GTTCTTAATTTGGTAGACTGTTGGAACAGAGAACTCAGAAAGTTTAAATATGCTTCTT 2299
Db 2197 GTTCTTAATTTGGTAGACTGTTGGAACAGAGAACTCAGAAAGTTTAAATATGCTTCTT 2256
Qy 2300 TGGATTAATGATGGGTCAATGCTTCAAGAAATGGAATCCAGTAGAAATGATACCAAG 2359
Db 2257 TGGATTAATGATGGGTCAATGCTTCAAGAAATGGAATCCAGTAGAAATGATACCAAG 2316
Qy 2360 ATGATCTTTGGTAAATTTAAATGCTGCTGATTCATCTCTCCCTGCTGAAAAATGAAGCCACTG 2419
Db 2317 ATGATCTTTGGTAAATTTAAATGCTGCTGATTCATCTCTCCCTGCTGAAAAATGAAGCCACTG 2376

QY	2420	GCCTGGCTTTGTTCTCTCCACCTCTTCTGCTCCAAATCAGAGGTCCATTGTTTCCAGTGGATG	2479
Db	2377	GCCTGGCTTTGTTCTCTCCACCTCTTCTGCTCCAGTCAGAGGTCCATTGTTTCCAGTGGATG	2436
QY	2480	CAAGAGGCCCATTTCTTGAGAGAGGAGGACCTCTTTTCCCCCCACCTCTCTCCACG-AGCCATG	2538
Db	2437	CAAGAGGCCCATTTCTTGAGAGAGGAGGACCTCTTTTCCCCCCACCTCTCTCCAGGAGGCCATG	2496
QY	2539	TTTGAGACTTCTCGAGATTAATTTTCCACAAAGG-ATTTCCAGGTCCACCACC-TGC	2594
Db	2497	TTTGAGACTTCTCGAGATTAATTTTCCACAGGGGGATTTCCAGGGTCCACCACCCTTGC	2556
QY	2595	TCCATTG-CAATCAGAAATGCTCTATCCACCGAGGGGTTTTCTCTCT-ACCT	2645
Db	2557	TCCATTGGCAATGAGGAATGCTTATCCACCAGGGGGTTTTCTCTCTTAACTT	2616
QY	2646	TCCCCAAGACCT--GGATTTTCCCCCAACCCCACTCT--GNAAGTAGAAGTGAG	2700
Db	2617	TCCCCAAGACNGGGGTTTTTCACAGGGGGATTTTTCACAGGGTCCAGGAGTGA	2676
QY	2701	TTCCCTCAGG-TTTGATTCACACTTCAAATGAGCCTGCTACTGAAATCCAGAACCA	2759
Db	2677	TTTCCCTTAGGTTTGATTTCAAATGAGCCTGCTACTGAAATCCAGAACCA	2736
QY	2760	GCAAG-AAACCTGACATATTTT-GCTCTCTTCAAAAGTAATTTGACTGATCTCATTT	2817
Db	2737	GCAAGAAACCTGACATATTTTGGGCTCTCTTCAAAGTAATTTGACTGATCTCATTT	2796
QY	2818	TCAGTTTAAAGTAACTGCTGTACTTAAAGTGATTAACATTTT-GCTCAAATGAGCTTAA	2876
Db	2797	TCAGTTTAAAGTAACTGCTGTACTTAAAGTGATTAACATTTTGGCTCAAATGAGCTTAA	2856
QY	2877	TGGAATATAATCTCAGGATGATTTTGTAAATAAGATGATTTAAATATGATCTTA	2936
Db	2857	TGGAATATAATCTCAGGATGATTTTGTAAATAAGATGATTTAAATATGATCTTA	2916
QY	2937	TCAGTAAATTTTCAATTTTATTTAGACGGTATACTATTTCATTTGATTAATCCAC	2996
Db	2917	TGAGTAAATTTTCAATTTTATTTAGACGGTATACTATTTCATTTGATTAATCCAC	2976
QY	2997	TATTATATAAACATAGTGGAGTTTATATATGATTAATCTTTCAGGTGGGAGGCTTAA	3056
Db	2977	TATTATATAAACATAGTGGAGTTTATATATGATTAATCTTTCAGGTGGGAGGCTTAA	3036
QY	3057	ATTCGAAAGTCTGCTCTTTATGCCAAGACTGTATTCTGTGGTTGTGGACAAATGTG	3116
Db	3037	ATTCGAAAGTCTGCTCTTTATGCCAAGACTGTATTCTGTGGTTGTGGACAAATGTG	3096
QY	3117	AAAGTAATCTTATGCTTAAATAAATATGTTGATTTAAAAA	3158
Db	3097	AAAGTAATCTTATGCTTAAATAAATATGTTGATTTAAAAA	3138
RESULT 4			
ABQ77424			
ID	ABQ77424 standard; cDNA; 2928 BP.		
XX	AC ABQ77424;		
XX	10-MAY-2003 (first entry)		
DT	Human CGDD cDNA 6715627CB1 SEQ ID 38.		
DE	Human; cell growth, differentiation and death; CGDD; cardiast; cancer;		
XX	cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;		
KW	neurological disorder; gene; ss.		
XX	Homo sapiens.		
OS	WO2003014322-A2.		
PN	20-FEB-2003.		
XX			
PD			
XX			

PF	08-AUG-2002; 2002WO-US025465.	
XX		
XX	08-AUG-2001; 2001US-0311017P.	
PR	17-AUG-2001; 2001US-0313070P.	
PR	17-AUG-2001; 2001US-0313071P.	
PR	24-AUG-2001; 2001US-0314678P.	
PR	31-AUG-2001; 2001US-0316692P.	
PR	07-SEP-2001; 2001US-0317913P.	
PR	14-SEP-2001; 2001US-0322182P.	
PR	07-DEC-2001; 2001US-0340747P.	
PR	20-DEC-2001; 2001US-0342761P.	
PR	29-MAR-2002; 2002US-0369129P.	
XX		
XX	(INCYTE) INCYTE GENOMICS INC.	
XX		
XX	Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;	
PI	Dugan BM, Elliott VS, Emerling BM, Forsythe IJ, Gietzen KJ;	
PI	Gorvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AUA, Ison CH;	
PI	Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;	
PI	Sprague MW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;	
PI	Yue H;	
XX		
XX	WPI; 2003-268197/26.	
XX	P-PSDB; ABG74685.	
XX		
XX	New polypeptide associated with cell growth, differentiation and death,	
PR	useful for preparing a composition for diagnosing or treating a disease	
PT	e.g., cardiovascular or neurological disorders or cancer.	
PT		
XX		
XX	Claim 93; Page 251-252; 267pp; English.	
XX		
XX	This invention describes a novel polypeptide associated with cell growth,	
CC	differentiation and death (CGPD) which has cardiant, cytostatic and	
CC	neuroprotective activity. The polypeptides of the invention are useful	
CC	for preparing a composition for diagnosing, or treating a disease or	
CC	condition associated with decreased expression or overexpression of	
CC	functional CGPD e.g., cardiovascular or neurological disorders or cancer.	
CC	The polynucleotides of the invention can be used for gene therapy.	
CC	ABQ77414-ABQ77441 encode the human CGPD polypeptides represented in	
CC	ABG74675-ABG74701	
XX		
XX	Sequence 2928 BP; 909 A; 597 C; 660 G; 762 T; 0 U; 0 Other;	
XX		
XX	Query Match 79.5%; Score 2515.2; DB 7; Length 2928;	
PI	Best Local Similarity 93.6%; Pred.No. 0;	
PI	Matches 2708; Conservative 0; Mismatches 158; Indels 28; Gaps 7	
QY	90 CTAAGCGGGCGGGCTCAGACCAAGCGCTCCCTCAGAGTGTAAGGTGAACAGAGGGCC 149	
DB	35 CTTGTGCGGGTCGGGCTCGGACCTCGCTCGCTCGGATGTAAGGTATTAACAGAGGGTC 94	
QY	150 AGGGAGGTGTGGGGGACACATGGGCCGTGTAGGCCTGTGGTCCCGCGTTCGCCAG 209	
DB	95 GGGATGGTCAGGTAGGCTGTGTAGGCGCTGT-----CGGTGCGCTGTCCCCATGC 148	
QY	210 CTCCTCCCGCAGCCCGCTCCACAGTGTCCGCTCCGCTGGTGTGTCACTGCGCATTCGG 269	
DB	149 TCCCTCCAGCAGCCGCTCCGAGTGTCTCCACTCCGCT----TGCCTGGTGGGATTCGG 204	
QY	270 GTTCAGACCCAGGCTCGGTTCCTCACCGCTTGTGT-----GGCCAGTGT 319	
DB	205 GTTCCGAGCCGAGGCTGTGTCTCGCGTTATTGTGGCCCCCAGCGCCGGGT 264	
QY	320 ACTGCGGTGACCGCAGAGCAGCTCCAGCGTATFGAGAGAGCTGTGTACCCCTCAGC 379	
DB	265 ACTGTGGCCACCACAGAGCAGCTTTCGCGCTATFGAGAGAGCCCGGGCTTACCCCTCAAC 324	
QY	380 CCTACTCGGGCTGTCTCGAGAGAGCTTACCGAGTGTGGCAGAGTACTACCTGAGAGTA 439	
DB	325 CGTATTGGGGCTGCTCTCTGAGAGAGCTACCGAGGTTGTGGCAGCACTGCTGAAGGTA 384	
QY	440 TGAGACCATGAGAACTTATGGTTTTCATCGGAAGTGTGTGGTATGTGAGCTGTTA 499	

385	TGAGCAGCAATCTAATCTTTATTTGGTTTTCCATGGGAATTTGGTGATATGTGCAGCTGTTG	444
500	TTGGATTTTWTGTTCTCCITTTTTTGTGGAGAAAGTTTATAGTCGGTTAGAGTCGAC	559
445	TTGGATTTTWTGCTGTTCTCTTTTTTTTGTGGAGAAAGTTTATAGTCGGTTAGAGTCGAC	504
560	TTTACGTGGGAAGAGAGCAAAAACCTTGGTCCAACGCTTTCTGACATAATTTGAAGAAAAAT	619
505	TTTATGTGGACGAGAGAAAAGCTTGCTTAATGCTTCTGCACTAATTTGAAGAAAAAA	564
620	GTAAACTATTTGAAAAAATTTAGCCTTTATCCAAAAAGAGTATGAAGGCTATGAAGTAGAGT	679
565	GTAAACTACTTTGAAAAAATTTAGCCTTTGTCAAAAAGAGATGAAGGCTATGAAGTAGAGT	624
680	CATCTTTTAGAGGATGCCAGCTTTGAGAAGGAGGCAGCAGAAGAAGCAGCAAGTTTCGAGG	739
625	CATCTTTAAGGATGCCAGCTTTGAGAAGGAGGCAAC--AGAAGCACAAAGTTTCGAGG	681
740	CAACCTGTGAAAAAGCTGAACAGTCCOATTTCTGAACTTGAGGATGAAATCCCTCTGCTAG	799
682	CAACCTGTGAAAAAGCTGAACAGTCCAAATTTCTGAACTTGAGGATGAAATCACTCTGCTAG	741
800	AAAAAGACTTTAAAAACAAGAGAAATCTAAACATTTCTCAACAAGATGAATGATGCCGATA	859
742	AAAAAGATTTAAAGAAGAGAAATCCAAACATTTCTGACAAGATGAATGATGCCGATA	801
860	TTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCAAAATAGCTG	919
802	TTTCAAAAAGGATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCAAAATAGCTG	861
920	AAGCCAAAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAAGCAGCGGCTTATAGCAATAA	979
862	AAGCCAAAAATGACCTTCAAGATATTTCAAATGAATGAAGAAGCAGCTGAAGATAGCAATAA	921
980	AGATGCTTTGAATGAAAAATTTCTCAACTTCAGACAAGCCATAACAGCTTTTTCAGCAAG	1039
922	AGATGCTTTGAATGAAAAATTTCTCAACTTCAGGAAGCCAGAACAGCTTTT--GCAAG	978
1040	AAGCTGAAGTATGAAAAGGAGAGTGAGTGAACCTTTAATAACAGAAAAATTAACATTTGAAG	1099
979	AAGCTGAAGTATGAAAAGAACAGTGAGTGAACCTTTAATAACAGAAAAATTAACATTTGAAG	1038
1100	ACTCCAAGTACAGCGAGAACAGTTCTGAATGATTAAGAAAAATCACATCAAGACCTGTA	1159
1039	ACTCCAAGTACATGCGAGAACAGTTCTGAATGATTAAGAAAAATCACATCAAGACCTGTA	1098
1160	CTGACACTTTGCCAATGATGAAGATCAGGCTGCTGTCTTTGAAGAGACACAAACGGATG	1219
1099	CTGAACGCTTGTAAAGATGAAGATTTGGGCTGCTATGCTTGGAGAAGACATAACCGATG	1158
1220	ATGATACCTGGAAATTAAGAGTGAACAGTGAATCGGAAAAATGTGCTTACTTAGATAATC	1279
1159	ATGATAACTTGGAAATTAAGAAATGAACAGTGAATCGGAAAAATGTGCTTACTTAGATAATC	1218
1280	CTCCAAAGAGAGCTTTGAAGAAAATCTGATTCATGCTGCTAAAGTTAAATGCTTTCTTAAAAA	1339
1219	CTCCAAAGAGAGCTTTGAAGAAAATCTGATTCATGCTGCTAAAGTTAAATGCTTTCTTAAAAA	1278
1340	CTTTAGAGGAGAGAAACCAAAATTTATATTCAGTTGTCTGAGTTGATTAACAAACAGG	1399
1279	CTTTAGAGGAGAGAAACCAAAATTTATATTCAGTTGTCTGAGTTGATTAACAAACAGG	1338
1400	AAGAGCTTTACAGACATATTTAAAAATCTTCAGACTCAACAGCATCTTTGCAAGTACAGAAA	1459
1339	AAGAGCTTTACAGACATATTTAAAAATCTTCAGACTCAACAGCATCTTTGCAAGTACAGAAA	1398
1460	ACACACATTTTGAATAATGAGAAATCAGAGCTTCAACGAACTTTAAAGTAAATGACTGAAT	1519
1399	ACACACATTTTGAATAATGAGAAATCAGAGCTTCAACGAACTTTAAAGTAAATGACTGAAT	1458
1520	TATATCAAGAAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGGAAAAATATCCGT	1579
1459	TATATCAAGAAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGGAAAAATATCCGT	1518

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:11:04 ; Search time 212 Seconds
(without alignments)
8282.380 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttaacatcatgaagag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A-COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5E-COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A-COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B-COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6C-COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/6D-COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	3164	100.0	3164	US-09-744-314-14
2	2416.2	76.4	2770	US-09-744-314-14
3	78	2.5	116592	US-09-744-314-14
4	70.4	2.2	7218	US-09-744-314-14
5	59.8	1.9	3095	US-09-744-314-14
6	56.2	1.8	11131	US-09-744-314-14
7	54.4	1.7	636	US-09-744-314-14
8	52.6	1.7	1664976	US-09-744-314-14
9	52.6	1.7	1664976	US-09-744-314-14
10	51.6	1.6	837	US-09-744-314-14
11	51.2	1.6	19124	US-09-744-314-14
12	51	1.6	1664976	US-09-744-314-14
13	50.8	1.6	2637	US-09-744-314-14
14	50	1.6	3222	US-09-744-314-14
15	49.2	1.6	6306	US-09-744-314-14
16	49	1.5	832	US-09-744-314-14
17	49	1.5	6326	US-09-744-314-14
18	49	1.5	168575	US-09-744-314-14
19	48.8	1.5	30549	US-09-744-314-14
20	48	1.5	580073	US-09-744-314-14
21	47.8	1.5	8607	US-09-744-314-14
22	47.4	1.5	665	US-09-744-314-14
23	47.4	1.5	6156	US-09-744-314-14
24	47	1.5	10640	US-09-744-314-14
25	46.8	1.5	2169	US-09-744-314-14
26	46.6	1.5	6866	US-09-744-314-14
27	46.4	1.5	9347	US-09-744-314-14

28	46.4	1.5	640681	4	US-09-750-988-1	Sequence 1, Appli
29	46.2	1.5	6306	4	US-10-204-708-50	Sequence 50, Appl
30	46.2	1.5	7664	4	US-10-204-708-84	Sequence 84, Appl
31	45.8	1.4	3446	4	US-09-392-714-13	Sequence 13, Appl
32	45.6	1.4	3636	4	US-09-134-001C-1983	Sequence 1983, Ap
33	45.2	1.4	11050	4	US-10-204-708-86	Sequence 86, Appl
34	45	1.4	2017	4	US-09-404-879A-72	Sequence 72, Appl
35	45	1.4	2017	4	US-09-338-933-72	Sequence 72, Appl
36	45	1.4	2017	4	US-09-215-681-72	Sequence 72, Appl
37	45	1.4	2017	4	US-09-215-003A-72	Sequence 72, Appl
38	44.4	1.4	633	4	US-09-134-001C-578	Sequence 578, App
39	44.4	1.4	1707	4	US-09-134-001C-931	Sequence 931, App
40	44.4	1.4	4766	5	US-09-134-001C-931	Sequence 10, Appl
41	44.4	1.4	5433	3	US-08-929-329-11	Sequence 11, Appl
42	44.2	1.4	6317	4	US-10-204-708-11	Sequence 11, Appl
43	44.2	1.4	19124	2	US-08-487-826B-13	Sequence 13, Appl
44	44.2	1.4	19233	4	US-10-204-708-46	Sequence 46, Appl
45	44.2	1.4	20674	4	US-09-641-638-651	Sequence 651, App

ALIGNMENTS

RESULT 1

US-09-976-594-320

Sequence 320, Application US/09976594

Patent No. 6673549

GENERAL INFORMATION:

APPLICANT: Furness, Michael

APPLICANT: Buchbinder, Jenny

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

FILE REFERENCE: PA-0041 US

CURRENT APPLICATION NUMBER: US/09/976,594

CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/240,409

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 1143

SOFTWARE: PERL Program

SEQ ID NO 320

LENGTH: 3164

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6673549 3768043CB1

US-09-976-594-320

Query Match 100.0%; Score 3164; DB 4; Length 3164;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTTATCATGAAGAAGACAGCAATTTTAAAGTGTTTACCAAAAATAAAGCTT 60

Db 1 AGGTTTATCATGAAGAAGACAGCAATTTTAAAGTGTTTACCAAAAATAAAGCTT 60

QY 61 CAAAATATGTGATGTGAAGTGCAGAACTGACAGGCGGGCGGCTCAGACCGGTGC 120

Db 61 CAAAATATGTGATGTGAAGTGCAGAACTGACAGGCGGGCGGCTCAGACCGGTGC 120

QY 121 CTCAGATTAAGTGTAAAGTGCAGGAGGCGGCGGAGTGTGGGGGACCAATGGCGCTG 180

Db 121 CTCAGATTAAGTGTAAAGTGCAGGAGGCGGCGGAGTGTGGGGGACCAATGGCGCTG 180

QY 181 TGAGGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCG 240

Db 181 TGAGGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCG 240

QY 241 CTCGCGTGTGGTGTACGTGGCGATTCGGTTCAGACCCAGGCTGGTGTCTCCACC 300

Db 241 CTCGCGTGTGGTGTACGTGGCGATTCGGTTCAGACCCAGGCTGGTGTCTCCACC 300

QY 301 GCTTGTGTGGCGAGTGTACTTGCCTGACCGCCAGCAGCCCTCGAGCTATGAGAG 360

Db 301 GCTTGTGTGGCGAGTGTACTTGCCTGACCGCCAGCAGCCCTCGAGCTATGAGAG 360

Claims 320
Plurality 7 DMW

Db 301 GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGAGCAGCTCGACGCTATGGAGGAG 360
Qy 361 CTTGTGTGTACCCCTACGCTACCTGGGCTGGTCTGGAGGAGCTACGAGAGTGTG 420
Db 361 CTTGTGTGTACCCCTACGCTACCTGGGCTGGTCTGGAGGAGCTACGAGAGTGTG 420
Qy 421 GCAGCACTACCTGAGAGTATGAGACAGATGAGAACTCTTATGGTTTCCATCGGAAC 480
Db 421 GCAGCACTACCTGAGAGTATGAGACAGATGAGAACTCTTATGGTTTCCATCGGAAC 480
Qy 481 GTGGTATGTGACGCTGTATTTGGAATTTTGTGTGTCTCTTCTTGTGTGAGAGTTT 540
Db 481 GTGGTATGTGACGCTGTATTTGGAATTTTGTGTGTCTCTTCTTGTGTGAGAGTTT 540
Qy 541 AGATCGGTGAGAGTGGCTTTACCTGGGAGAGAGCAAAACCTTGGTGCAACGCTTCT 600
Db 541 AGATCGGTGAGAGTGGCTTTACCTGGGAGAGAGCAAAACCTTGGTGCAACGCTTCT 600
Qy 601 GCACCTAATTTGAAGAAATGTAATCTTGAATTTTGGTCTTCTTCTTCTTCTTCTTCT 660
Db 601 GCACCTAATTTGAAGAAATGTAATCTTGAATTTTGGTCTTCTTCTTCTTCTTCTTCT 660
Qy 661 GAAGCTATGAAGTATGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCA 720
Db 661 GAAGCTATGAAGTATGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCA 720
Qy 721 GAAGCAGAGGTTTGGAGGCACTGTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAA 780
Db 721 GAAGCAGAGGTTTGGAGGCACTGTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAA 780
Qy 781 GATGAATCTCTGTCTAGAAAAGACTTAAACCAAGAGAAATCTTCAACAA 840
Db 781 GATGAATCTCTGTCTAGAAAAGACTTAAACCAAGAGAAATCTTCAACAA 840
Qy 841 GATGAATCTGAGGCGGATTTTCAAAAGATATACAGTCTTGAAGAGTGAATCTTCA 900
Db 841 GATGAATCTGAGGCGGATTTTCAAAAGATATACAGTCTTGAAGAGTGAATCTTCA 900
Qy 901 CTCAAATCAAAATAGCTGAAGCCTTCTGCAAGCAATTTAAATGAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAAGCCTTCTGCAAGCAATTTAAATGAGTGAAGAA 960
Qy 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTCAAGAGCCAT 1020
Db 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTCAAGAGCCAT 1020
Qy 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATA 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATA 1080
Qy 1081 CAGAAATAAATTTGAAGACTTCAAAAGTACACGAGAACTGAAGTTCTGAATGATAAG 1140
Db 1081 CAGAAATAAATTTGAAGACTTCAAAAGTACACGAGAACTGAAGTTCTGAATGATAAG 1140
Qy 1141 AATCATCATCAAGCCTGACTGACACTGTCCTGCAATGATCAAGATCAGGCTGCTG 1200
Db 1141 AATCATCATCAAGCCTGACTGACACTGTCCTGCAATGATCAAGATCAGGCTGCTG 1200
Qy 1201 GAAGAGACACACCGGATGATGAATCTGGAATTTAGAGTGAACAGTGAATTCGGA 1260
Db 1201 GAAGAGACACACCGGATGATGAATCTGGAATTTAGAGTGAACAGTGAATTCGGA 1260
Qy 1261 GGTGCTTACTAGATTAATCTTCAAAAGAGCTTTGAAGAACTGATTCATGCTCTAAG 1320
Db 1261 GGTGCTTACTAGATTAATCTTCAAAAGAGCTTTGAAGAACTGATTCATGCTCTAAG 1320
Qy 1321 TTAATGCTTCTTTAAAAACCTTAGAAGGAGAAAGAAACCAATTTATATTCAAGTGTCT 1380
Db 1321 TTAATGCTTCTTTAAAAACCTTAGAAGGAGAAAGAAACCAATTTATATTCAAGTGTCT 1380
Qy 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATATTAATAATCTTCAAGCTCAACAA 1440
Db 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATATTAATAATCTTCAAGCTCAACAA 1440

Qy 1441 GCATCTTTTCAGTCAAGAAACACACATTTTGAATTCAGAACTCAAGAGCTTCAACAGAAA 1500
Db 1441 GCATCTTTTCAGTCAAGAAACACACATTTTGAATTCAGAACTCAAGAGCTTCAACAGAAA 1500
Qy 1501 CTTAAAGTAAATGACTGAATTTATCAAGAAATGAAATGAAATCTCCACAGGAAATTAACA 1560
Db 1501 CTTAAAGTAAATGACTGAATTTATCAAGAAATGAAATGAAATCTCCACAGGAAATTAACA 1560
Qy 1561 GTAGAGGAAATTTATCGGTTTAGAGAAAGAGAAACCTTTCTTAAAGTAGATGAAAGATC 1620
Db 1561 GTAGAGGAAATTTATCGGTTTAGAGAAAGAGAAACCTTTCTTAAAGTAGATGAAAGATC 1620
Qy 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA 1680
Db 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA 1680
Qy 1681 TTGGAGAGAACTATTCTTCTTATCAAGGCGAGATTTATTTCCATGAGAAAGAGCAGAT 1740
Db 1681 TTGGAGAGAACTATTCTTCTTATCAAGGCGAGATTTATTTCCATGAGAAAGAGCAGAT 1740
Qy 1741 GATAATTTGGTTGSCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
Db 1741 GATAATTTGGTTGSCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
Qy 1801 GCTCAACACAGACAAATTTAACTGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAGAT 1860
Db 1801 GCTCAACACAGACAAATTTAACTGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAGAT 1860
Qy 1861 CTTTATGCACTCGATGTTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
Db 1861 CTTTATGCACTCGATGTTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
Qy 1921 TCACCAATTTGGTTGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Db 1921 TCACCAATTTGGTTGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Qy 1981 GAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGCAGGC 2040
Db 1981 GAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGCAGGC 2040
Qy 2041 CAGAGGAACTCTTGACACCATCAGATACGATGAAAGAGGAGAACTCAAGCTGTGATAG 2100
Db 2041 CAGAGGAACTCTTGACACCATCAGATACGATGAAAGAGGAGAACTCAAGCTGTGATAG 2100
Qy 2101 TTAACGATCTCTATAGGCTCCTCTGACACTGAGCTTCTGTCACTCCATCCATGGGACCA 2160
Db 2101 TTAACGATCTCTATAGGCTCCTCTGACACTGAGCTTCTGTCACTCCATCCATGGGACCA 2160
Qy 2161 GACCGTAGGATGATTTTCTCTGCGCAGGACAAATCATATCTGATTCAGCCCTTCTCTCA 2220
Db 2161 GACCGTAGGATGATTTTCTCTGCGCAGGACAAATCATATCTGATTCAGCCCTTCTCTCA 2220
Qy 2221 CAAAGCAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCCAGACAACTCAGA 2280
Db 2221 CAAAGCAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCCAGACAACTCAGA 2280
Qy 2281 AGTTTAAATATGCTTCTTTGGATAAAATGATGGGTCAATGCCCTTCAAGAAATGGAATCC 2340
Db 2281 AGTTTAAATATGCTTCTTTGGATAAAATGATGGGTCAATGCCCTTCAAGAAATGGAATCC 2340
Qy 2341 AGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCT 2400
Db 2341 AGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCT 2400
Qy 2401 GCTGAAATGAAGCCACTGSCCTTGTCTTCTCCACCTTCTCTCCATTCAGAGGT 2460
Db 2401 GCTGAAATGAAGCCACTGSCCTTGTCTTCTCCACCTTCTCTCCATTCAGAGGT 2460
Qy 2461 CCATGTTCCAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 2461 CCATGTTCCAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520

QY 2521 CCTCCTCAGGAGCCATGTTTGGAGCTTCTCGAGATATATTTTCCACCAGGGAATTTCCCA 2580
 Db |||||
 QY 2521 CCTCCTCAGGAGCCATGTTTGGAGCTTCTCGAGATATATTTTCCACCAGGGAATTTCCCA 2580
 Db |||||
 QY 2581 GGTTCACACACCTGCTCCATTTTCCAAATGAGAAATGCTATCCACCGAGGGGTTTTCCCTCT 2640
 Db |||||
 QY 2581 GGTTCACACACCTGCTCCATTTTCCAAATGAGAAATGCTATCCACCGAGGGGTTTTCCCTCT 2640
 Db |||||
 QY 2641 TACCTTCCCCCAGGAGCTGGATTTTCCCCCACCACCCACATCTCGAAGGTAGAGTGAG 2700
 Db |||||
 QY 2641 TACCTTCCCCCAGGAGCTGGATTTTCCCCCACCACCCACATCTCGAAGGTAGAGTGAG 2700
 Db |||||
 QY 2701 TTCCCTCAGGTTGATTCACCTTTCAAAATGAGCTGCTACTGAACTCCAGAACACAG 2760
 Db |||||
 QY 2701 TTCCCTCAGGTTGATTCACCTTTCAAAATGAGCTGCTACTGAACTCCAGAACACAG 2760
 Db |||||
 QY 2761 CAAGAAACCTGCAATATTTTTCCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA 2820
 Db |||||
 QY 2761 CAAGAAACCTGCAATATTTTTCCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA 2820
 Db |||||
 QY 2821 GTTTAAGTAACCTGCTTACTTAAGTGATTAACATTTTCTCABATTTGAAGCTTAATGGA 2880
 Db |||||
 QY 2821 GTTTAAGTAACCTGCTTACTTAAGTGATTAACATTTTCTCABATTTGAAGCTTAATGGA 2880
 Db |||||
 QY 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTTAAATATGAATCTTATGAG 2940
 Db |||||
 QY 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTTAAATATGAATCTTATGAG 2940
 Db |||||
 QY 2941 TAAATTTTCAATTTTATTTAGAGGTAATCACTATTTCAATTTGATTAATCCACTATT 3000
 Db |||||
 QY 2941 TAAATTTTCAATTTTATTTAGAGGTAATCACTATTTCAATTTGATTAATCCACTATT 3000
 Db |||||
 QY 3001 ATATAAACCAATAGTGGAGTTTATATATGTAATCTTTTCCAGTGGGGAGGCTTTAAATTC 3060
 Db |||||
 QY 3001 ATATAAACCAATAGTGGAGTTTATATATGTAATCTTTTCCAGTGGGGAGGCTTTAAATTC 3060
 Db |||||
 QY 3061 TGAAGTCTGCTTATGCAAGAACTGATTTTACTGTTTGGACAAATGTGAAG 3120
 Db |||||
 QY 3061 TGAAGTCTGCTTATGCAAGAACTGATTTTACTGTTTGGACAAATGTGAAG 3120
 Db |||||
 QY 3121 TAACTTTATGCTTAAATAAATATAGTGTATTTTAAATAAAAAA 3164
 Db |||||
 QY 3121 TAACTTTATGCTTAAATAAATATAGTGTATTTTAAATAAAAAA 3164
 Db |||||

RESULT 2

US-09-976-594-322
 ; Sequence 322, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchbinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 322
 ; LENGTH: 2770
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 039946.19
 US-09-976-594-322

Query Match 76.4%; Score 2416.2; DB 4; Length 2770;
 Best Local Similarity 92.3%; Pred. No. 0;
 Matches 2631; Conservative 0; Mismatches 83; Indels 135; Gaps 3;

QY 310 GCCAGTGTACTGCGGTGACCCGACAGAGAGCCTCGACGCTATGAGGAGCCTGGTGT 369
 Db |||||
 QY 370 ACCCTCAGCCCTACCTGCGGTGCTTCCCTGGAGAGCTACGACAGAGTTGTGCGACACTA 429
 Db |||||
 QY 370 ACCCTCAGCCCTACCTGCGGTGCTTCCCTGGAGAGCTACGACAGAGTTGTGCGACACTA 429
 Db |||||
 QY 430 CTTGAGAGTATGAGACCAAGATGAGATCTTATGTTTCCATCGGAATCGTGGTATGT 489
 Db |||||
 QY 430 CTTGAGAGTATGAGACCAAGATGAGATCTTATGTTTCCATCGGAATCGTGGTATGT 489
 Db |||||
 QY 490 GCAGCTGTTATGGAATTTTGTGTTCTCTCTTTTGTGGAGAGTTTATAGATCGTT 549
 Db |||||
 QY 490 GCAGCTGTTATGGAATTTTGTGTTCTCTCTTTTGTGGAGAGTTTATAGATCGTT 549
 Db |||||
 QY 550 AGAGTCGGCTTTACGTGGGAGAGAGCAAAATCTTGGTGCACGCTTCTGACACTAAT 609
 Db |||||
 QY 550 AGAGTCGGCTTTACGTGGGAGAGAGCAAAATCTTGGTGCACGCTTCTGACACTAAT 609
 Db |||||
 QY 610 GAAGAAATGTAATCTACTTTGAAATTTTAGCTTTATTCAAAAGAGTATGAAGGCTAT 669
 Db |||||
 QY 610 GAAGAAATGTAATCTACTTTGAAATTTTAGCTTTATTCAAAAGAGTATGAAGGCTAT 669
 Db |||||
 QY 670 GAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGCAGAGACGCA 729
 Db |||||
 QY 670 GAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGCAGAGACGCA 729
 Db |||||
 QY 730 AGTTTGGAGCAACCTCTGAAAGAGTGAACAGCTTCCAAATCTGAACTTGAAGTGAATC 789
 Db |||||
 QY 730 AGTTTGGAGCAACCTCTGAAAGAGTGAACAGCTTCCAAATCTGAACTTGAAGTGAATC 789
 Db |||||
 QY 790 CTCTGCTAGAGAAAGAGTTTAAACAAAGAGAAATCTTAAACATTTCTCAAGATGAATG 849
 Db |||||
 QY 790 CTCTGCTAGAGAAAGAGTTTAAACAAAGAGAAATCTTAAACATTTCTCAAGATGAATG 849
 Db |||||
 QY 850 ATGCGGATATTTCAAAAGATGATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCA 909
 Db |||||
 QY 850 ATGCGGATATTTCAAAAGATGATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCA 909
 Db |||||
 QY 910 CAAATAGCTGAGCCAAATCACTGCAAGACATTTTAAATAGTGAAGAACGCGGCT 969
 Db |||||
 QY 910 CAAATAGCTGAGCCAAATCACTGCAAGACATTTTAAATAGTGAAGAACGCGGCT 969
 Db |||||
 QY 970 ATAGCAATTAAGATGCTTTTGAATGAAATTTCAACTTCAGACCAAGCCTTAAACAGCTT 1029
 Db |||||
 QY 970 ATAGCAATTAAGATGCTTTTGAATGAAATTTCAACTTCAGACCAAGCCTTAAACAGCTT 1029
 Db |||||
 QY 1030 TTTGAGCAAGAGCTGAAATGATGAAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1089
 Db |||||
 QY 1030 TTTGAGCAAGAGCTGAAATGATGAAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1089
 Db |||||
 QY 1090 ACATTTGAGAGTCCAAAGTACACGACAGCAAGTCTGATGATGAAGAAATCAATC 1149
 Db |||||
 QY 1090 ACATTTGAGAGTCCAAAGTACACGACAGCAAGTCTGATGATGAAGAAATCAATC 1149
 Db |||||
 QY 1150 AAGACCTTGACTGGACACTTCCCAATGATGAAAGATCAGGCTGCTGCTTTGAGAGAGAC 1209
 Db |||||
 QY 1150 AAGACCTTGACTGGACACTTCCCAATGATGAAAGATCAGGCTGCTGCTTTGAGAGAGAC 1209
 Db |||||
 QY 1210 ACAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
 Db |||||
 QY 1210 ACAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
 Db |||||
 QY 1270 TTAGATATATCTCCAAAGAGGAGCTTTGAGAGAACTGATTCATGCTGCTTAAATGCT 1329
 Db |||||
 QY 1270 TTAGATATATCTCCAAAGAGGAGCTTTGAGAGAACTGATTCATGCTGCTTAAATGCT 1329
 Db |||||
 QY 1330 TCTTTAAAGAGCTTAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
 Db |||||
 QY 1330 TCTTTAAAGAGCTTAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
 Db |||||
 QY 1390 ABAACAGGAGAGCTTACAGAGCATATTTAAATATTTTCAAGCTTCAACAGAGCATTTTG 1449
 Db |||||
 QY 1390 ABAACAGGAGAGCTTACAGAGCATATTTAAATATTTTCAAGCTTCAACAGAGCATTTTG 1449
 Db |||||

1085 AAAACAAAGGAGAGCTTACAGACATATTAATAATCTTCAGACTCAACAGCATCTTTG 1144
1450 CAGTCAGAAAAACACACATTTTGAATAAGAGATCAGAAAGCTTCAACAGAACTTAAAGTA 1509
1145 CAGTCAGAAAAACACACATTTTGAATAAGAGATCAGAAAGCTTCAACAGAACTTAAAGTA 1204
1510 ATGACTGAATTATATCAAGAAAAATGAATGAATCTCCACAGAAATTAACAGTAGAGGAA 1569
1205 ATGACTGAATTATATCAAGAAAAATGAATGAATCTCCACAGAAATTAACAGTAGAGGAA 1264
1570 AATTATCGGTAGAGAAAGAGAGAACTTTCTTAAAGTAGATGAAGATCAGCCATGCC 1629
1265 AATTATCGGTAGAGAAAGAGAGAACTTTCTTAAAGTAGATGAAGATCAGCCATGCC 1324
1630 ACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAAATTCAGAGAGA 1689
1325 ACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAAATTCAGAGAGA 1384
1690 ACTATTCTATTCTTCAAGGGCAGATTAATTTCCCATGAGAAAAAGACACATGATAATGG 1749
1385 ACTATTCTATTCTTCAAGGGCAGATTAATTTCCCATGAGAAAAAGACACATGATAATGG 1444
1750 TTGGCAGCTCGGAATCTCAAGAGAACTCAATGATTTAAGGAAAGAAATGCTCACAAC 1809
1445 TTGGCAGCTCGGAATCTCAAGAGAACTCAATGATTTAAGGAAAGAAATGCTCACAAC 1504
1810 AGACAAAAATTAACCTGAACAGAGCTTAAATTTGAATTTAGAAAAAGATCCTTATGCA 1869
1505 AGACAAAAATTAACCTGAACAGAGCTTAAATTTGAATTTAGAAAAAGATCCTTATGCA 1564
1870 CTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCAATTTCCCATATGGTCCCTCACCATTG 1929
1565 CTCGATGTTCCAAATACAGCAATTTGGC ----- 1591
1930 GGTTGGCCTTCATCTGAACAAAGAGCTTTTCTCTCTCTCCTCCAACCTTTGTTGGAGGGTCCA 1989
1592 ----- 1591
1990 CTCAGACTCTCACCCTTGTCTTCAGGGGAGAGAGAGGCTCAGAGGCCAGGGAAAT 2049
1592 ----- AGAGGCTCAGAGGCCAGGGAAAT 1615
2050 CCTCTGGACCATCAGATTACCAATGAAGAGAGAGATCAAGCTGTGATAGGTTAACCGAT 2109
1616 CCTCTGGACCATCAGATTACCAATGAAGAGAGAGATCAAGCTGTGATAGGTTAACCGAT 1675
2110 CCTCATAGGGCTCCCTCTGACACTGGGTCTGTGTGACCTCCATGGGACAGAGCCGTAGG 2169
1676 CCTCATAGGGCTCCCTCTGACACTGGGTCTGTGTGACCTCCATGGGACAGAGCCGTAGG 1735
2170 ATGATGTTTCTCCGCGCAGACAATCATATCTGATTCAGCCCTTCCCTCCACAAAGGCAA 2229
1736 ATGATGTTTCTCCGCGCAGACAATCATATCTGATTCAGCCCTTCCCTCCACAAAGGCAA 1795
2230 GACAGATTTGTTCTTAATCTGTAGACTGTCTGGACACAGAGAACTCAGAAAGTTTAAAT 2289
1796 GACAGATTTGTTCTTAATCTGTAGACTGTCTGGACACAGAGAACTCAGAAAGTTTAAAT 1855
2290 ATGCTTCTTTGATATAATGGATGGGTCAATGCTTCAGAAATGAATCCAGTAGAAT 2349
1856 ATGCTTCTTTGATATAATGGATGGGTCAATGCTTCAGAAATGAATCCAGTAGAAT 1915
2350 GATACCAAGAGATGATCTTGATAATTTAAATGTGCTGATTCTCTCCCTGCTGAAAT 2409
1916 GATACCAAGAGATGATCTTGATAATTTAAATGTGCTGATTCTCTCCCTGCTGAAAT 1975
2410 GAGCCACTGGCTGGCTTGTCTCCACTCTGCTCCATCAGAGGTCCATGTTT 2469
1976 GAGCCACTGGCTGGCTTGTCTCCACTCTGCTCCATCAGAGGTCCATGTTT 2035
2470 CCAGTGGATCAAGAGGCCCATCTTTGAGAGAGAGACCTCTTCCCTCCCACTCTCTCCA 2529

2036 CCAGTGGATCAAGAGGCCCATCTTTGAGAGAGAGACCTCTTCCCTCCCACTCTCTCCA 2095
2530 GGAGCCATCTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGATTTCCAGGTCCACCA 2589
2096 GGAGCCATCTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGATTTCCAGGTCCACCA 2155
2590 CTTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTCCCTTACCTTCC 2649
2156 CTTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTCCCTTACCTTCC 2215
2650 CCAAGACCTGGATTTTCCCTCCCACTCTGAAAGGTAGAAAGTGAAGTGAAGTGAAGTGA 2709
2216 CCAAGACCTGGATTTTCCCTCCCACTCTGAAAGGTAGAAAGTGAAGTGAAGTGAAGTGA 2275
2710 GGTTCGATTTCCACCTTTCAAAATGAGGCTGCTACTGAAACATCCAGAACCCAGCAAGAAC 2769
2276 GGTTCGATTTCCACCTTTCAAAATGAGGCTGCTACTGAAACATCCAGAACCCAGCAAGAAC 2335
2770 TGAACATATTTTGTCTCTCTCAAAAGTAAATTTGACCTGATCTCATTTTTCAGTTTAAAGTA 2829
2336 TGAACATATTTTGTCTCTCTCAAAAGTAAATTTGACCTGATCTCATTTTTCAGTTTAAAGTA 2395
2830 ACTGCTGTTACTTAACTGATTACACTTTTCTCAAAATGAAGCTTAAATGAAGTGAAGTGA 2889
2396 ACTGCTGTTACTTAACTGATTACACTTTTCTCAAAATGAAGCTTAAATGAAGTGAAGTGA 2455
2890 CTGAGATAGTATTTTGAATTAAGATGATTTAAATATGAATCTTATGAGTAAATTTAT 2949
2456 CTGAGATAGTATTTTGAATTAAGATGATTTAAATATGAATCTTATGAGTAAATTTAT 2515
2950 TCAATTTATTTTGAAGCTGATTAATTTCAATTTTGAATTTAAATCCTATTATATAAACA 3009
2516 TCAATTTATTTTGAAGCTGATTAATTTCAATTTTGAATTTAAATCCTATTATATAAACA 2575
3010 ATAGTGGAGTTTATATATGATTAATCTTTCAGTGGGAGGCTTTAAATCTTGAAGTCTG 3069
2576 ATAGTGGAGTTTATATATGATTAATCTTTCAGTGGGAGGCTTTAAATCTTGAAGTCTG 2635
3070 TGTCTTTATCCCAAGAACTGATTTTACTGTGGTGTGGCAAAATGTGAAAGTAACTTTAT 3129
2636 TGTCTTTATCCCAAGAACTGATTTTACTGTGGTGTGGCAAAATGTGAAAGTAACTTTAT 2695
3130 GCTTAAATAAATATATAGTTGATTTAAAAA 3158
2696 GCTTAAATAAATATATAGTTGATTTAAAAA 2724

RESULT 3
US-09-818-512-3/c
; Sequence 3, Application US/09818512
; Patent No. 6537780
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0601192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A, T, C or G
US-09-818-512-3

Query Match 2.5%; Score 78; DB 4; Length 116592;
Best Local Similarity 57.3%; Pred. No. 2.6e-10;

Matches 141; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 1644 GACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAAATGGAGAGAACTATTCTTTA 1703
Db 42013 GACCTCAAGAAGATGGCCCAAGATCTGGCCAGGAATGGAGAAAGCGCTTCCAATGC 41954
QY 1704 TCAAGGCGAGATTATTCCTATGAGAAAGACACATGATAATTGGTTGGCGCTCGGAA 1763
Db 41953 TTGAGAAAGCTCTCTTCTATGAGAAAGACAGCCCAAGATGCTGGTGGCTTCCAGTC 41894
QY 1764 TGCTGAAGAAACCTCAATGATTAAAGAAAGAAATGCTCACAACAGACAAAATTAAC 1823
Db 41893 CACTGTGCGAGTGTCTAGTGTCTGTGAAAGAAATGACCAATAGGCAGAGCCAGT 41834
QY 1824 TGAACAGAGCTTAAATTTCACTTTTAGAAAGAGATCTTATGCATCGATGTTCCAAA 1883
Db 41833 TGACTGTGAGTCAAGTTTCAGCCTTCCCAAGCTGTCTTGGTCCAGGATTTCCGTC 41774
QY 1884 TACAGC 1889
Db 41773 TGCAGC 41768

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHIEFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZcgt-F1s
US-08-232-463-14

Query Match

2.2%; Score 70.4; DB 1; Length 7218;

Best Local Similarity 13.6%; Pred. No. 5.2e-09;
Matches 74; Conservative 233; Mismatches 239; Indels 0; Gaps 0;
QY 1241 TCAACAGTGAATCGGAAATAGTGTCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAGA 1300
Db 1570 TAAATATGAAAGTGGTATTGTGACTAGCGTAGCATCGCTTCTAGAGCATCTATTCCAGT 1511
QY 1301 AACTGATTTCATGCTGCTAAAGTTAAATGCTTCTTAAACCTTAGAAGGAGAAAGAAC 1360
Db 1510 TTCAAAAACCGCATAGGATCATCTGTAATTAACCTATCTATCAAGTAGTTAAAGAGA 1451
QY 1361 AAATTTATTTTCAAGTGTCTGAGTTGATAAAACAAAGGAGAGCTTACAGAGCATATTA 1420
Db 1450 TAGAAGATTTGGTACRR 1391
QY 1421 AAATCTTCAGACTCAACAAGCATCTTTGCAGTCAGAAAACACACATTTTGAATATGAGA 1480
Db 1390 RRR 1331
QY 1481 ATCAGAGCTTCAACAGAACTTAAAGTAATGACTGAAATTAATATCAAGAAAATGAATGA 1540
Db 1330 RRR 1271
QY 1541 AACTCCACAGAAATTAACAGTAGAGGAAATTAATCGTTAGAGAAAGAGAAACTTT 1600
Db 1270 RRR 1211
QY 1601 CTAAAGTAGATGAAAGATCAGCCATGCACTGAGAGCTGAGACCTATAGAAAGCGAG 1660
Db 1210 RRR 1151
QY 1661 CCAAGATCTTGAAGAAATGAGAGAACTATTCATCTTATCAAGGGCAGATATTT 1720
Db 1150 RRR 1091
QY 1721 CCATGAGAAAGACACATGATAATGTTGGCAGCTCGGAATGCTGCAAGAAACCTCA 1780
Db 1090 RRR 1031
QY 1781 ATGATT 1786
Db 1030 TTAATT 1025

RESULT 5

5231168-1
; Patent No. 5231168
; APPLICANT: DZIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409.658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 1;
; LENGTH: 3095
5231168-1

Query Match 1.9%; Score 59.8; DB 6; Length 3095;
Best Local Similarity 44.3%; Pred. No. 2.9e-06;
Matches 290; Conservative 0; Mismatches 362; Indels 3; Gaps 1;
QY 1199 TTGAAGAGACACAAACGGATGATGATAACCTGGAATTAGAAGTGAACAGTGAATCGGAAA 1258
Db 965 TAGATCATAAACAGTTGATCCAGAAATAGTAGAAGTTGAAGAAATCTCTTCAGAACTAC 1024
QY 1259 ATGGTGCTTACTAGATAATCTCCAAAGAGCTTTGAAGAACTGATTCATGCTGTA 1318
Db 1025 ATGAAAATGAAGTGGCTCATCCAGAAATTTGTTGAAATTTAGGAAGTTTCTCTGAACCA 1084
QY 1319 AGTTAAATGCTTCTTTAAACCTTTAGAGGAGAAAGAACCCAAATTTATATTCAGTTGT 1378
Db 1085 ATCAAAATAACGAATTCAGAAATTAATGAGATGATAAAGTGCACATATTC--GC 1141

QY 1379 CTGAAGTTGATATAAACAAGAGAGAGCTTACAGACATATTAATAATCTTCAGACTCAAC 1438
DB 1142 ATGAATAGTAGAGTAGAGAGAAATACCTTCAGAGAGATGATATAAATGAAAAAGTTGAC 1201
QY 1439 AAGCATCTTTGAGTCAGAGAAACACACATTTTGAATAATGAGAAATCAGAGCTTCAACAGA 1498
DB 1202 ATGAATAGTAGAGTAGAGAGAAATCTTACCAGAGAGATATAAATGAAAAAGGTCACATG 1261
QY 1499 AACTTAAATGATGACTGAAATATATATCAAGAAATGAAATGAAATCCACAGGAAATTA 1558
DB 1262 AATAGTAGAGGTTGAGAAATCTTACCAGAGAGATGATATAAATGAAAAAGTTGACATG 1321
QY 1559 CAGTAGAGAAATATATCGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAGAAAGA 1618
DB 1322 AATAGTAGAGGTTGAGAAATCTTACCAGAGAGATATAAATGAAAAAGGTCACATGAA 1381
QY 1619 TCAGCCATCCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAG 1678
DB 1382 TAGTAGAGGTTGAGAAATCTTACCAGAGAGATATAAATGAAAAAGTTGACATGAAATAG 1441
QY 1679 AATGGAGAGAACTATTCATCTTATCAAGGCGAGATTTTCCCATGAGAGAAAAAGCAC 1738
DB 1442 TAGAAGTTGAAGAAATCTTACCAGAGAGATATAAATGAAAAAGGTCACATGAAATAGTAG 1501
QY 1739 ATGATAATTTGGTGGCAGCTCGAATGCTGAGAGAACTCAATGATTTAGGAGAGAA 1798
DB 1502 AGTTGAGAGAAATCTTACCAGAGAGATATAAATGAAAAAGTTTCAATGAAATAGTAGAG 1561
QY 1799 ATGCTCACACAGACAAAAATTAACCTGAACAGAGCTTAAATTTGAACCTTTTGA 1853
DB 1562 TTGAAGAAATCTTACCAGAGAGATATAAATGAAAAAGGTCACATGAAATAGTAGA 1616

RESULT 6

US-10-204-708-28
; Sequence 28, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR FILING DATE: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; LENGTH: 11131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-28

Query Match 1.8%; Score 56.2; DB 4; Length 11131;
Best Local Similarity 46.3%; Pred. No. 7.8e-05;
Matches 187; Conservative 0; Mismatches 218; Indels 0; Gaps 0;
QY 2744 AACATCCAGAACACACAGCAGAGAACTGACAAATATTTTCTCTCTTCAAAAGTAAATTT 2803
DB 10465 AAAAATTAAGAGAAAGAAATTAAGAGTAGTAGTATTTATTTTATTTAGAGTTGAATATT 10524

QY 2804 GACTGATCTCATTTTCAGTTTAAAGTAACCTGCTGCTTAACTTAAGTATTACACCTTTTGCTCA 2863
DB 10525 TAAGGAABATAGGTATGGTTTATTGATTTATTATTATTAGATAAATATAAATTTTGAAGT 10584
QY 2864 AATTGAAGCTTAATGAATTAATATTCACAGATAGTATTTTGTAAATAAGATGATTTTA 2923
DB 10585 TGGGTATGTTTATGAGTTTATAAGTAGTATATATTTTATTTTATTTTAAATATATTG 10644
QY 2924 AATATGAATCTTATGAGTAAATTTATTTCAATTTTATTTTAGACGCTATATTTCAAT 2983
DB 10645 AATTTTAAAAAATTAATTTGTTTTTATTTTATGATGAATTTTGAAGTATATTAAT 10704
QY 2984 TTGATTAATCCACTATATATAACATAGTGGAGTTTATATATGTAATCTTTCAGGT 3043
DB 10705 AGTTATTTATTTATGAGATTTATATGTTTAGATTTTTTTTTTGTAGTATTTTGCCT 10764
QY 3044 GGGGAGCTTTAAATTTCTGAAGTCTGCTTTATGCCAAGAACTGTTATTTACTGTGGTT 3103
DB 10765 GTATTATAATTTTATTTATATAATTTTGTGAGATAAGTTTGTATTGTTTTTATTTT 10824
QY 3104 GTGACAAATGTGAAGTAACTTTATGCTTAAATAAATATATAGTT 3148
DB 10825 ATTGAAGAAATTTGAGAAGTAGAGTGTAAAGAAATTTGTTTTTAGGT 10869

RESULT 7

US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalleis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: RF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: April 21, 2004, 02:53:20 ; Search time 1286 Seconds
 (without alignments)
 11092.765 Million cell updates/sec

Title: US-09-744-314-14
 Perfect score: 3164
 Sequence: 1 aggttaaccatgaagaag.....agttgattataaaaaaaa 3164

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
 13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2.*
 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
 17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
 18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
 19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2414.6	76.3	2814	9	US-09-880-107-3419
2	2414.6	76.3	2814	9	Sequence 3419, App
3	2414.6	76.3	2814	13	US-09-967-768A-191
4	2414.6	76.3	2814	13	Sequence 191, App
5	2414.6	76.3	2814	15	US-10-081-056-335
6	2414.6	76.3	2814	15	Sequence 335, App
7	2414.6	76.3	2814	15	US-10-305-654-335
8	2414.6	76.3	2814	15	Sequence 335, App
9	2414.6	76.3	2814	15	US-10-223-083-335
10	2414.6	76.3	2814	15	Sequence 335, App
11	2414.6	76.3	2814	15	US-10-223-087-335
12	2414.6	76.3	2814	15	Sequence 335, App
13	2414.6	76.3	2814	15	US-10-223-089-335
14	2397.2	75.8	2602	16	US-10-223-081-335
					Sequence 335, App
					Sequence 27, Appl

15	2303.4	72.8	2614	16	US-10-161-927-29	Sequence 29, Appl
16	2159.8	68.3	2351	9	US-09-864-761-21534	Sequence 21534, A
c 17	2148.6	67.9	2334	9	US-09-864-761-30648	Sequence 30648, A
18	2111.4	66.7	2588	16	US-10-161-493-53	Sequence 53, Appl
19	2036.2	64.0	2483	16	US-10-161-493-55	Sequence 55, Appl
20	1954.6	61.8	4292	9	US-09-729-674-41	Sequence 41, Appl
21	1838.8	58.1	1986	9	US-09-864-761-4795	Sequence 4795, App
c 22	1797.8	56.8	1961	9	US-09-864-761-14030	Sequence 14030, A
23	912.8	28.8	925	15	US-10-029-386-20575	Sequence 20575, A
c 24	871.8	27.6	883	15	US-10-029-386-20733	Sequence 20733, A
25	777.2	24.6	1044	9	US-09-822-830A-596	Sequence 596, App
c 26	565.8	17.9	569	15	US-10-029-386-6964	Sequence 6964, App
c 27	562.6	17.8	569	15	US-10-029-386-7023	Sequence 7023, App
c 28	559.8	17.7	643	9	US-09-920-300A-163	Sequence 163, App
c 29	559.8	17.7	643	14	US-10-033-528-163	Sequence 163, App
c 30	559.8	17.7	643	15	US-10-099-926-163	Sequence 163, App
c 31	479.8	15.2	559	9	US-09-864-761-12319	Sequence 12319, A
32	466	14.7	597	13	US-10-027-632-265229	Sequence 265229, A
33	466	14.7	597	16	US-10-027-632-265229	Sequence 265229, A
34	464.4	14.7	597	13	US-10-027-632-265228	Sequence 265228, A
35	464.4	14.7	597	16	US-10-027-632-265228	Sequence 265228, A
c 36	462	14.6	593	13	US-10-027-632-264782	Sequence 264782, A
c 37	462	14.6	593	16	US-10-027-632-264782	Sequence 264782, A
c 38	460.4	14.6	593	13	US-10-027-632-264783	Sequence 264783, A
c 39	460.4	14.6	593	16	US-10-027-632-264783	Sequence 264783, A
c 40	458.8	14.5	543	15	US-10-060-036-1501	Sequence 1501, App
c 41	438.4	13.9	440	9	US-09-998-598-1014	Sequence 1014, App
c 42	436.4	13.8	480	13	US-10-027-632-268158	Sequence 268158, A
c 43	436.4	13.8	480	16	US-10-027-632-268158	Sequence 268158, A
c 44	434.8	13.7	480	13	US-10-027-632-268157	Sequence 268157, A
c 45	434.8	13.7	480	13	US-10-027-632-268159	Sequence 268159, A

ALIGNMENTS

RESULT 1

US-09-880-107-3419
 ; Sequence 3419, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3419
 ; LENGTH: 2814
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U73682
 US-09-880-107-3419

Query Match 76.3%; Score 2414.6; DB 9; Length 2814;
 Best Local Similarity 92.3%; Pred. No. 0;
 Matches 2630; Conservative 0; Mismatches 84; Indels 135; Gaps 3;
 QY 310 GCCCAGTGTACTGGGTGACCGCAGACGACCTCGCGTATGGAGAGCGCTGTGCT 369
 Db 5 GCCCGGGGTACTGTGGCGACACACAGAGACAGCTTTGGCGCTATGGAGAGCGCGGGCT 64
 QY 370 ACCCTCAGCCCTACTGGGTGCTGCTGGAGAGCTACCGAGAGTGTGGCAGCACTA 429

Db 65 ACCCTCAACCGTATTTGGGGTCTCTCGAGGAGCTACGAGGGTTTGGCCACACTG 124
QY 430 CTGTGAGATGATGAGACCACTGATGAGATCTTATGTTTTCATCGGAATCGTGTGTATGT 489
Db 125 CTGGAAGTATGAGACCACTGATCTATCTTTATGTTTTCATCGGAATCGTGTATGT 184
QY 490 GAGCTGTTATTTGGATTTTGTGTTTCTCTTTTGTGAGAGTGTGATGATGATGATGATG 549
Db 185 GAGCTGTTGTTGGATTTTGTGTTTCTCTTTTGTGAGAGTGTGATGATGATGATGATG 244
QY 550 AGGAGTCGGCTTTACCTGGGAAGAGAGCAAACTTTGGTGCACGCTTTCTGGACTAATT 609
Db 245 AGGAGTCGGCTTTATGTTGGAGCGAGAGAAAAGCTTCTCTATGCTTTCTGGACTAATT 304
QY 610 GAAGAAAATGTAACCTCTGAAAATTTAGCCCTTATCAAAAAGATGATGAGGCTAT 669
Db 305 GAAGAAAATGTAACCTCTGAAAATTTAGCCCTTATCAAAAAGATGATGAGGCTAT 364
QY 670 GAAGTAGAGTCTATCTTTAGAGATGCCAGCTTTGAGAGGAGGAGAGAGAGACAGA 729
Db 365 GAAGTAGAGTCTATCTTTAAGATGCCAGCTTTGAGAGGAGGAGAGAGAGACAGA 421
QY 730 AGTTGGAGCAACCTGTGAAAAGCTGACAGGTCCAAATCTGAACTTTGAGAGTGAATC 789
Db 422 AGTTGGAGCAACCTGTGAAAAGCTGACAGGTCCAAATCTGAACTTTGAGAGTGAATC 481
QY 790 CTCTGTCTAGAAAAGACTTTAAAAGAGAGAACTCTAAACATCTCTCAACAGATGAATTG 849
Db 482 CTCTGTCTAGAAAAGAGATTTAAAAGAGAGAACTCTAAACATCTCTCAACAGATGAATTG 541
QY 850 ATGGGGATPATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCA 909
Db 542 ATGGGGATPATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCA 601
QY 910 CAATAGCTGAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAGAGAGAGAGAG 969
Db 602 CAATAGCTGAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAGAGAGAGAGAG 661
QY 970 ATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGCAAGCCATAAAACAGCTT 1029
Db 662 ATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGCAAGCCATAAAACAGCTT 721
QY 1030 TTTACAGCAAGAGCTGAGTATGGAAGAGAGAGAGTGAAGTGAATTAATAACAGAAATA 1089
Db 722 TT---GCAAGAGCTGAGATGGAAGAGAGAGAGTGAAGTGAATTAATAACAGAAATA 778
QY 1090 ACATTTGAAGAGCTCCAAAGTACAGCAGAAACAGTCTTGAATGATAAAGAAAATCACATC 1149
Db 779 ACATTTGAAGAGCTCCAAAGTACAGCAGAAACAGTCTTGAATGATAAAGAAAGTCAATC 838
QY 1150 AAGACCTGACTGGACACTTGCATGATGAAGATCAGGCTGCTGTGTTGAAGAGAGAC 1209
Db 839 AAGACTCTGACTGAACGCTTGTAAAGATGAAGATTTGGGCTGCTATGCTTTGAAGAGAC 898
QY 1210 ACAACGGATGATGATAAAGCTGGAATTAGAAGTGAACAGTGAATCGGAAAATGGTCTTAC 1269
Db 899 ATAACGGATGATGATAAAGCTGGAATTAGAAGTGAACAGTGAATCGGAAAATGGTCTTAC 958
QY 1270 TTAGATAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGCTGAAGTAAATGCT 1329
Db 959 TTAGATAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGCTGAAGTAAATGCT 1018
QY 1330 TCTTTAAAACCTTTAGAAG 1389
Db 1019 TCTTTAAAACCTTTAGAAG 1078
QY 1390 AAAACAAAAGAGAGCTTTACAGAGCATATTAATAATCTTACAGCTCAACAGAGCTTTTG 1449
Db 1079 AAAACAAAAGAGAGCTTTACAGAGCATATTAATAATCTTACAGCTCAACAGAGCTTTTG 1138
QY 1450 CAGTCAGAAAACACATTTGGAATGAGATCAGAGCTTCAACAGAGAGAGAGAGAGAGAGAG 1509
Db 1139 CAGTCAGAAAACACATTTGGAATGAGATCAGAGCTTCAACAGAGAGAGAGAGAGAGAGAG 1198

QY 1510 ATGACTGAATTTATCAAGAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGGAA 1569
Db 1199 ATGACTGAATTTATCAAGAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGGAA 1258
QY 1570 AATTTCGGTTAGAGAAAAGAGAACTTTCTAAAGTAGATGAAGAGATCAGCCATGCC 1629
Db 1259 AATTTCGGTTAGAGAAAAGAGAACTTTCTAAAGTAGATGAAGAGATCAGCCATGCC 1318
QY 1630 ACTGAAGAGCTGAGAGACCTATAGAAAAGCAGGCAAAAGATCTTGAAGAAATTTGGAGAGA 1689
Db 1319 ACTGAAGAGCTGAGAGACCTATAGAAAAGCAGGCAAAAGATCTTGAAGAAATTTGGAGAGA 1378
QY 1690 ACTATTCTATTTATCAAGGGCAGATTTTCCATGAGAAAAGACACATGATTAATTCG 1749
Db 1379 ACTATTCTATTTATCAAGGGCAGATTTTCCATGAGAAAAGACACATGATTAATTCG 1438
QY 1750 TTGGCAGCTCGGAATGCTGAAAAGAACTCAATGATTTAAGGAAAAGAAAATGCTCAACAC 1809
Db 1439 TTGGCAGCTCGGAATGCTGAAAAGAACTCAATGATTTAAGGAAAAGAAAATGCTCAACAC 1498
QY 1810 AGACAAAATTAATCTGAAAAGACAGCTTTAAATTTGAACTTTTGAAGAAAAGATCTTATGCA 1869
Db 1499 AGACAAAATTAATCTGAAAAGACAGCTTTAAATTTGAACTTTTGAAGAAAAGATCTTATGCA 1558
QY 1870 CTGATGTTCCAAATACAGCATTTGGCAGAGAGACATTCGCCATATGCTCCCTCACCATTTG 1929
Db 1559 CTGATGTTCCAAATACAGCATTTGGC----- 1585
QY 1930 GGTGCGCTTCATCTGAAAAGACAGAGCTTTTCTCTCTCCAACTTTGTTGGAGGTTCCA 1989
Db 1586 ----- 1585
QY 1990 CTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGAGAGGCTCACAGGCCCCAGGGAAT 2049
Db 1586 -----AGAGGCTCACAGGCCCCAGGGAAT 1609
QY 2050 CCTCTGACCATCAGATTAACAAATGAAGAGAGAGAAATCAAGCTGTGATAGTTAAACCGAT 2109
Db 1610 CCTCTGACCATCAGATTAACAAATGAAGAGAGAGAAATCAAGCTGTGATAGTTAAACCGAT 1669
QY 2110 CCTCATAGGCTCCCTCTGACACTGGCTCTCTGACCTCCATGGACAGGAGCCGATAGG 2169
Db 1670 CCTCATAGGCTCCCTCTGACACTGGCTCTCTGACCTCCATGGACAGGAGCCGATAGG 1729
QY 2170 ATGATGTTTCTCCGCGCAGGACAAATCATATCTCTGATTCAGCCCTTCTCCACAAAGGCAA 2229
Db 1730 ATGATGTTTCTCCGCGCAGGACAAATCATATCTCTGATTCAGCCCTTCTCCACAAAGGCAA 1789
QY 2230 GACAGATTTTGTCTAATCTGCTAGACTGTCTGGACAGGAGAACTCAGAGTTTAAAT 2289
Db 1790 GACAGATTTTGTCTAATCTGCTAGACTGTCTGGACAGGAGAACTCAGAGTTTAAAT 1849
QY 2290 ATGCTTTCTTTGGATAAAATGGATGGGTCAATGCTCTCAGAAATGGAATCCAGTAGAAAT 2349
Db 1850 ATGCTTTCTTTGGATAAAATGGATGGGTCAATGCTCTCAGAAATGGAATCCAGTAGAAAT 1909
QY 2350 GATACCAAGAGATGATCTTGTAAATTTAATGCTGCTGATTCATCTCTCCCTGCTGAAAT 2409
Db 1910 GATACCAAGAGATGATCTTGTAAATTTAATGCTGCTGATTCATCTCTCCCTGCTGAAAT 1969
QY 2410 GAAGCCACTGCGCTTGTCTTCTCCACTCTTCTGCTCCAAATCAGAGTCCATTTGTTT 2469
Db 1970 GAAGCCACTGCGCTTGTCTTCTCCACTCTTCTGCTCCAAATCAGAGTCCATTTGTTT 2029
QY 2470 CCAGTGAATCAGAGGCCATTTCTTGAAGAGGAGCTCTTCTCCCTCCCTCCCTCCCA 2529
Db 2030 CCAGTGAATCAGAGGCCATTTCTTGAAGAGGAGCTCTTCTCCCTCCCTCCCTCCCA 2089
QY 2530 GGAGCCATGTTGGAGCTTCTCGAGATTTATTTCCACAGGAGATTTCCAGGTCACCA 2589
Db 2090 GGAGCCATGTTGGAGCTTCTCGAGATTTATTTCCACAGGAGATTTCCAGGTCACCA 2149

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 06:17:26 ; Search time 9919 Seconds
(without alignments)
11223.801 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttcaatccatgaagaag.....agtgatttaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA_Main.*

- 1: /cgn2_6/ptodata/2/pna/PCTUS COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US105 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6054 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Db	841		GATGAATTGATGGCGGATATTTTCAAAAGATATACAGTCTCTAGAAATGAGTCAAAATCC	900
Qy	901	CTCAAAATCACAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA	960	
Db	901	CTCAAAATCACAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA	960	
Qy	961	CGACGGCTATAGCAATAAAAGATGCTTTGAATGAAAAATTTCTCAACTTCAGACAAGCCAT	1020	
Db	961	CGACGGCTATAGCAATAAAAGATGCTTTGAATGAAAAATTTCTCAACTTCAGACAAGCCAT	1020	
Qy	1021	AAACAGCTTTTTCAGACAAGAGCTGAAGTATGAAAGGAGAGTGAAGTGAATTAATAAA	1080	
Db	1021	AAACAGCTTTTTCAGACAAGAGCTGAAGTATGAAAGGAGAGTGAAGTGAATTAATAAA	1080	
Qy	1081	CAGAAAATAACATTTGAAGACTCCAAAGTACACGCGAGAACAAAGTTCTGAATGATAAGAA	1140	
Db	1081	CAGAAAATAACATTTGAAGACTCCAAAGTACACGCGAGAACAAAGTTCTGAATGATAAGAA	1140	
Qy	1141	AATCAGATCAAGACCCCTGACTGACACTTGCCTCAATGATGAAGATCAGCTGCTGCTT	1200	
Db	1141	AATCAGATCAAGACCCCTGACTGACACTTGCCTCAATGATGAAGATCAGCTGCTGCTT	1200	
Qy	1201	GAAGAAAGACACACGGATGATTAACCTCGAAATTAGAAGTGAACAGTGAATCGGAAAAAT	1260	
Db	1201	GAAGAAAGACACACGGATGATTAACCTCGAAATTAGAAGTGAACAGTGAATCGGAAAAAT	1260	
Qy	1261	GGTGCCTTACTTAGATAATCCTCCAAAGGAGCTTTTGAGNAACTGATTCATGCTGCTAAG	1320	
Db	1261	GGTGCCTTACTTAGATAATCCTCCAAAGGAGCTTTTGAGNAACTGATTCATGCTGCTAAG	1320	
Qy	1321	TTAAATGCTTCTTTTAAAAACCTTAGAAGGAGAAAGAAACCAAATTTATATTCAAGTTGTCT	1380	
Db	1321	TTAAATGCTTCTTTTAAAAACCTTAGAAGGAGAAAGAAACCAAATTTATATTCAAGTTGTCT	1380	
Qy	1381	GAAGTTTGATAAAAACAAAGGAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	1440	
Db	1381	GAAGTTTGATAAAAACAAAGGAGAGCTTACAGAGCATATTAAAAATCTTCAGACTCAACAA	1440	
Qy	1441	GCATCTTTGCAGTCAGAAAAACACATTTTGAAGAAATGAGAACTCAGAACTTCAACAGAAA	1500	
Db	1441	GCATCTTTGCAGTCAGAAAAACACATTTTGAAGAAATGAGAACTCAGAACTTCAACAGAAA	1500	
Qy	1501	CTTAAAGTAATGACTGAATTATATCAAGAAAAATGAAATGAAATCCACAGAAATTAACA	1560	
Db	1501	CTTAAAGTAATGACTGAATTATATCAAGAAAAATGAAATGAAATCCACAGAAATTAACA	1560	
Qy	1561	GTAGAGAAAAATATTCGGTTAGAGAAAGAGAAACTTTCTAAAGTAGTGAAGAGATC	1620	
Db	1561	GTAGAGAAAAATATTCGGTTAGAGAAAGAGAGAAACTTTCTAAAGTAGTGAAGAGATC	1620	
Qy	1621	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGAAGAGAA	1680	
Db	1621	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGAGCCAAAGATCTTTGAAGAGAA	1680	
Qy	1681	TTGAGAGAACTATTCAATCTTATCAAGGGCAGATTATTTCCCATGAGAAAAAGCAAT	1740	
Db	1681	TTGAGAGAACTATTCAATCTTATCAAGGGCAGATTATTTCCCATGAGAAAAAGCAAT	1740	
Qy	1741	GATAATTTGGTTGGCAGCTCGGAAATGCTGAAAGAAACCTCAATGATTTAAGCAAGAAAAAT	1800	
Db	1741	GATAATTTGGTTGGCAGCTCGGAAATGCTGAAAGAAACCTCAATGATTTAAGCAAGAAAAAT	1800	
Qy	1801	GCTCACACAGACAAAAATTAACCTGAAACACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1860	
Db	1801	GCTCACACAGACAAAAATTAACCTGAAACACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1860	
Qy	1861	CCTTATGCACCTCGATGTTCCAAATACAGCATTTTGGCAGAGAGCATTCGCCATATGTCCTC	1920	
Db	1861	CCTTATGCACCTCGATGTTCCAAATACAGCATTTTGGCAGAGAGCATTCGCCATATGTCCTC	1920	
Qy	1921	TCACCATTTGGTTGGCCTTCACTCTGAAACAAAGAGCTTTTCTCTCTCTCCCACTTTTGTG	1980	

D	b	1921	T C A C C A T T G G G T T G G C C T T C A T C T G A A A C A G A G C T T T T T C T C T C C T C C A A C T T T G T T G 1988
Q	y	1981	G A G G T C C A C T C A G A C T C T C A C C T T T T G C T T C C A G G G G A G G A A G A G A G C T C A C A G G C 2040
D	b	1981	G A G G T C C A C T C A G A C T C T C A C C T T T G C T T C C A G G G G A G A A G A G A G C T C A C A G G C 2040
Q	y	2041	C A G G G A A T C C T C T G A C C A T C A G A T T A C C A A T G A A G A G G A A A T C A A G C T G T G A T A G G 2100
D	b	2041	C C A G G G A A T C C T C T G G A C C A T C A G A T T A C C A A T G A A A G A G G A A N T C A A G C T G T G A T A G G 2100
Q	y	2101	T T A A C C G A T C C T C A T A G G G T C C C T C T G A C A C T F G G G T C T C T G T C A C C T C A T C G G G A C C A G 2160
D	b	2101	T T A A C C G A T C C T C A T A G G G T C C C T C T G A C A C T F G G G T C T C T G T C A C C T C A T C G G A C C A G 2160
Q	y	2161	G A C C G T A G A T G A T T T C C T C C G C A G G A C A A T C A T A T C C T G A T T C A G C C C T T C T C C A 2220
D	b	2161	G A C C G T A G A T G A T T T C C T C C G C A G G A C A A T C A T A T C C T G A T T C A G C C C T T C T C C A 2220
Q	y	2221	C A A A G G C A A C A G A G A T T T G C T T C T A A T T C T G T P A G A C T G T C T G A C A G C A G A G A A C T C A G A 2280
D	b	2221	C A A A G G C A A C A G A G A T T T G C T T C T A A T T C T G T P A G A C T G T C T G A C A G C A G A G A A C T C A G A 2280
Q	y	2281	A G T T T T A A T A T G C T T C T T T P G S A A A A T G G A T G G G T C A A T G C C T C A A T G C C T C A G A A A T G S A A T C C 2340
D	b	2281	A G T T T T A A T A T G C T T C T T T P G S A A A A T G G A T G G G T C A A T G C C T C A A T G C C T C A G A A A T G S A A T C C 2340
Q	y	2341	A G T A G A A A T G A T A C C A A A G A T G A T C T T G T P A A T T T A A A T G T G C T C A T T C A T C T C T C C C T 2400
D	b	2341	A G T A G A A A T G A T A C C A A A G A T A T C T T T G T P A A T T T A A A T G T G C T G A T T C A T C T C T C C C T 2400
Q	y	2401	G C T G A A A A T G A A G C A C T G C C C T G G C T T T G T T C C T C C A C C T C T T G C T C A A T C A G A G G T 2460
D	b	2401	G C T G A A A A T G A A G C A C T G C C C T G G C T T T G T T C C T C C A C C T C T T G T C C A A T C A G A G G T 2460
Q	y	2461	C C A T T G T T C C A G T G G A T G C A A G A G C C C A T T C T T G A R A G A G A C C T C C T T T C C C C C C A 2520
D	b	2461	C C A T T G T T C C A G T G G A T G C A A G A G C C C A T T C T T G A R A G A G A C C T C C T T T C C C C C C A 2520
Q	y	2521	C C T C C T C C A G A G A C C A T G T T T G G A G C T T C T C G A G A T T A T T T T C C A C A A G G G A T T T C C C A 2580
D	b	2521	C C T C C T C C A G A G A C C A T G T T T G G A G C T T C T C G A G A T T A T T T T C C A C A A G G G A T T T C C C A 2580
Q	y	2581	G G T C C A C C A C T G C T C A T T T G C A T G A G A A A T G T C T A T C A C G A G G G G T T T T C C T C C T 2640
D	b	2581	G G T C C A C C A C T G C T C A T T T G C A T G A G A A A T G T C T A T C A C G A G G G G T T T T C C T C C T 2640
Q	y	2641	T A C C T T C C C C A C A G A C T G G A T T T T C C C C C A C C C C C A C A T T C T G A A G G T A G A A G T A G 2700
D	b	2641	T A C C T T C C C C A C A G A C T G G A T T T T C C C C C A C C C C C A C A T T C T G A A G G T A G A A G T A G 2700
Q	y	2701	T T C C C C T C A G G T T T G A T T C C A C C T T C A A A T G A G C C T G C T A C T G A A C A T C C A G A A C C A C A G 2760
D	b	2701	T T C C C C T C A G G T T T G A T T C C A C C T T C A A A T G A G C C T G C T A C T G A A C A T C C A G A A C C A C A G 2760
Q	y	2761	C A G A A A C C T G A C A A T A T T T T G C T C T T C A A A G P A A T T T T G A C T G A T C T C A T T T T C A 2820
D	b	2761	C A G A A A C C T G A C A A T A T T T T G C T C T T C A A A G P A A T T T T G A C T G A T C T C A T T T T C A 2820
Q	y	2821	G T T T A A G T A A C T G C T G T A C T T A A G T A T A C A C T T T T G C T C A A A T T G A A G C T T A A T G G A 2880
D	b	2821	G T T T A A G T A A C T G C T G T A C T T A A G T A T A C A C T T T T G C T C A A A T T G A A G C T T A A T G G A 2880
Q	y	2881	A T T A T A A T T C C A G A T A G A T A T T T T G T A A A T A A G A T G A T T T A A A T G A A T G A A T C T T A T G A 2940
D	b	2881	A T T A T A A T T C C A G A T A G A T A T T T T G T A A A T A A G A T G A T T T A A A T G A A T G A A T C T T A T G A 2940
Q	y	2941	T A A A T A T T T C A A T T T A T T T T A G A C G G T A F A C T A T T T C A A T T T G A A T T A A T C C A C T A T T 3000
D	b	2941	T A B A T T A T T C A A T T T A T T T T A G A C G G T A F A C T A T T T C A A T T T G A A T T T A A T C C A C T A T T 3000
Q	y	3001	A T A T A A A C A A T A G T G G A G T T T T A T A T A T G T A A T C T T T C A G G T G G G A G C G C T T T A A A T T C 3060
D	b	3001	A T A T A A A C A A T A G T G G A G T T T T A T A T A T G T A A T C T T T C A G G T G G G A G C G C T T T A A A T T C 3060

QY 3061 TGAAGCTCTGCTTTATCCAGAGTGTATTTACTGTGGTGGCAAAATGTGAAG 3120
Db 3061 TGAAGCTCTGCTTTATCCAGAGTGTATTTACTGTGGTGGCAAAATGTGAAG 3120
QY 3121 TAACTTTATGCTTAAATAATATAGTGTATTTAAAAA 3164
Db 3121 TAACTTTATGCTTAAATAATATAGTGTATTTAAAAA 3164

RESULT 2
US-09-127-665-6
; Sequence 6, Application US/09127665
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Azimzai, Yalda
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 US
; CURRENT APPLICATION NUMBER: US/09/127,665
; CURRENT FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 3758043, BRSTN0124
US-09-127-665-6

Query Match 100.0%; Score 3164; DB 15; Length 3164;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAATAAGCTT 60
Db 1 AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCACCAAAAATAAGCTT 60
QY 61 CAAAATATGTGATGTGAAAACCTGCCAGAACTAAGCGCGCGGGCTCAGACCAGCGCTGC 120
Db 61 CAAAATATGTGATGTGAAAACCTGCCAGAACTAAGCGCGCGGGCTCAGACCAGCGCTGC 120
QY 121 CTCAGAGTAAAGTGTAAACAAGAGGCGAGGGAGGTGGTGGGAGCAACAATGGGCGTG 180
Db 121 CTCAGAGTAAAGTGTAAACAAGAGGCGAGGGAGGTGGTGGGAGCAACAATGGGCGTG 180
QY 181 TGAGGCGCTGGGTGCGCGCGTTCCTCCAGCTCCCGCGAGCGCTCCACAGTGGTCCG 240
Db 181 TGAGGCGCTGGGTGCGCGCGTTCCTCCAGCTCCCGCGAGCGCTCCACAGTGGTCCG 240
QY 241 CTCGGTGGTGTGTCACGTGGGATTCGGGTTCAGACCCAGGCTGGGTTCCTCCACC 300
Db 241 CTCGGTGGTGTGTCACGTGGGATTCGGGTTCAGACCCAGGCTGGGTTCCTCCACC 300
QY 301 GCTGTGTGGCCAGGTGTACTGCGGTGACCGCAGAGCAGCTCGAGCTATGAGGAG 360
Db 301 GCTGTGTGGCCAGGTGTACTGCGGTGACCGCAGAGCAGCTCGAGCTATGAGGAG 360
QY 361 CTTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGCTACGAGAGTTGTG 420
Db 361 CTTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGCTACGAGAGTTGTG 420
QY 421 CGAGCACTACCTGAGAGTATGAGCAGATGAGATCCTTATGTTTCCATCGGAAGT 480
Db 421 CGAGCACTACCTGAGAGTATGAGCAGATGAGATCCTTATGTTTCCATCGGAAGT 480
QY 481 GTGGTATGCGAGCTGTATTTGTTTGTCTCTCTTTTGTGGAGAGTTT 540
Db 481 GTGGTATGCGAGCTGTATTTGTTTGTCTCTCTTTTGTGGAGAGTTT 540

481 GTGGTATGCGAGCTGTATTTGTTTGTCTCTCTTTTGTGGAGAGTTT 540
541 AGATCGGTTAGGAGTGGCTTTTACGTGGAGAGAGCAAAATCTGGTGGCAAGCTTTCT 600
541 AGATCGGTTAGGAGTGGCTTTTACGTGGAGAGAGCAAAATCTGGTGGCAAGCTTTCT 600
601 GGACTAATTCAGAGAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAGAAAGAGTAT 660
601 GGACTAATTCAGAGAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAGAAAGAGTAT 660
661 GAAGGCTATGAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGGAGAGAA 720
661 GAAGGCTATGAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGGAGAGAA 720
721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAGAGCTGAAACAGCTCCAAATCTGAATCTGAG 780
721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAGAGCTGAAACAGCTCCAAATCTGAGCTTGA 780
781 GATGAATTCCTCTGTCTAGAAAAGACTTTAAACCAAGAGAAATCTAAACATTTCTCAACAA 840
781 GATGAATTCCTCTGTCTAGAAAAGACTTTAAACCAAGAGAAATCTAAACATTTCTCAACAA 840
841 GATGAATTCCTGTCTGTGCGGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
841 GATGAATTCCTGTCTGTGCGGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
901 CTCAAATCAGAAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATAGTGAAGAA 960
901 CTCAAATCAGAAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATAGTGAAGAA 960
961 CGACGGCTATAGCAATTAAGATGCTTTGAATGAATAATCTCAACTTCAGAGCAAGCCAT 1020
961 CGACGGCTATAGCAATTAAGATGCTTTGAATGAATAATCTCAACTTCAGAGCAAGCCAT 1020
1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
1081 CAGAAATACATTTGAGAGCTCCAAAGTACACGAGACAGCTTCTGAATGAATAAAGAA 1140
1081 CAGAAATACATTTGAGAGCTCCAAAGTACACGAGACAGCTTCTGAATGAATAAAGAA 1140
1141 AATCACATCAAGACCTCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGCTT 1200
1141 AATCACATCAAGACCTCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGCTT 1200
1201 GAAGAGACACAGGATGATGATTAACCTGGATTAAGAGTGAACAGTGAATCGGAAT 1260
1201 GAAGAGACACAGGATGATGATTAACCTGGATTAAGAGTGAACAGTGAATCGGAAT 1260
1261 GGTGCTTACTTAGATAATCTCCAAAGAGGAGCTTTGAAGAAACTGATTCATGCTGCTAAG 1320
1261 GGTGCTTACTTAGATAATCTCCAAAGAGGAGCTTTGAAGAAACTGATTCATGCTGCTAAG 1320
1321 TTAATGCTCTTTTAAACCTTAGAGAGGAGAGAAACCAATTTATTTAGTGTCT 1380
1321 TTAATGCTCTTTTAAACCTTAGAGAGGAGAGAAACCAATTTATTTAGTGTCT 1380
1381 GAAGTTCATAAAACAAAGAGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1440
1381 GAAGTTCATAAAACAAAGAGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1440
1441 GCATCTTTGAGTCAAGAAAACACATATTTGAAAATGAGANTCAGAGCTTCAACAGAAA 1500
1441 GCATCTTTGAGTCAAGAAAACACATATTTGAAAATGAGANTCAGAGCTTCAACAGAAA 1500
1501 CTTAAAGTATGATGATTAATTAAGAAAATGAATGAATGAATGAATGAATGAATGAATGAAT 1560
1501 CTTAAAGTATGATGATTAATTAAGAAAATGAATGAATGAATGAATGAATGAATGAATGAAT 1560
1561 GTAGAGAGAAATATTGCTTTAGAGAGAGAGAGAAATTTCTAAAGTAGATGAAGATC 1620
1561 GTAGAGAGAAATATTGCTTTAGAGAGAGAGAGAAATTTCTAAAGTAGATGAAGATC 1620

QY	1621	AGCATGCCACTGAAGAGCTGGAGACTTAGAAAGCGAGCCAAAATCTTTGAAGAAGAA	1680
DB	1621	AGCATGCCACTGAAGAGCTGGAGACTTAGAAAGCGAGCCAAAATCTTTGAAGAAGAA	1680
QY	1681	TTGGAGAGAAGCTATTCAATCTTCAATCAAGGCGAGATATTTCCCATGAGAAAAAGCACAT	1740
DB	1681	TTGGAGAGAAGCTATTCAATCTTCAATCAAGGCGAGATATTTCCCATGAGAAAAAGCACAT	1740
QY	1741	GATAAATGGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTAAAGAAAGAAAAT	1800
DB	1741	GATAAATGGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTAAAGAAAGAAAAT	1800
QY	1801	GCTCACAAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGCAACTTTTAGAAAAAGAT	1860
DB	1801	GCTCACAAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGCAACTTTTAGAAAAAGAT	1860
QY	1861	CCATTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCCATATGFPCC	1920
DB	1861	CCATTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCCATATGFPCC	1920
QY	1921	TCACCAATGGTTGGCCCTTCATCTGAAACAGAGCTTTCTCTCTCTCCCAACTTTGTTG	1980
DB	1921	TCACCAATGGTTGGCCCTTCATCTGAAACAGAGCTTTCTCTCTCTCCCAACTTTGTTG	1980
QY	1981	GAGGGTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGAGAGCTCACGAGC	2040
DB	1981	GAGGGTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGAGAGCTCACGAGC	2040
QY	2041	CCAGGGAATCCTCTGAGACCATCAGATTACCAATGARAAGAGAGAACTCAAGCTCTGATAGG	2100
DB	2041	CCAGGGAATCCTCTGAGACCATCAGATTACCAATGARAAGAGAGAACTCAAGCTCTGATAGG	2100
QY	2101	TTAAACCATCCTCATAGGGTCCCTCTGACACTGGGTCTCTGTACCTCCATCGGAGCCAG	2160
DB	2101	TTAAACCATCCTCATAGGGTCCCTCTGACACTGGGTCTCTGTACCTCCATCGGAGCCAG	2160
QY	2161	GACCGTAGGATGATTTTCTCCGCGAGGACATCATATCCTGATTCAGGCCCTTCCTCCA	2220
DB	2161	GACCGTAGGATGATTTTCTCCGCGAGGACATCATATCCTGATTCAGGCCCTTCCTCCA	2220
QY	2221	CAAAAGCAAGACAGATTTGTCTCTAAATCTCGTAGACTGTCTGACAGCAGAGAACTCAGA	2280
DB	2221	CAAAAGCAAGACAGATTTGTCTCTAAATCTCGTAGACTGTCTGACAGCAGAGAACTCAGA	2280
QY	2281	AGTTTTAATATGCCCTCTTTGGATAAATAGATGGGTCAATGCTTCAGAAATGGAAATCC	2340
DB	2281	AGTTTTAATATGCCCTCTTTGGATAAATAGATGGGTCAATGCTTCAGAAATGGAAATCC	2340
QY	2341	AGTAGAAATGATACCAAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT	2400
DB	2341	AGTAGAAATGATACCAAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT	2400
QY	2401	GCTGAAAATGAAGGCATGCGCCCTGGCTTTGTTCTCTCGACCTCTTGCTCCAAATCAGAGGT	2460
DB	2401	GCTGAAAATGAAGGCATGCGCCCTGGCTTTGTTCTCTCGACCTCTTGCTCCAAATCAGAGGT	2460
QY	2461	CCATTGTTTCCAGTGGATGCAAGAGGCCATTCTTGAGAAAGAGACCTCTTTCCCCCCCA	2520
DB	2461	CCATTGTTTCCAGTGGATGCAAGAGGCCATTCTTGAGAAAGAGACCTCTTTCCCCCCCA	2520
QY	2521	CCTCCTCCAGAGGCCATGTTTGGAGCTTCTGAGATTATTTTCCACCAAGGGATTTCCCA	2580
DB	2521	CCTCCTCCAGAGGCCATGTTTGGAGCTTCTGAGATTATTTTCCACCAAGGGATTTCCCA	2580
QY	2581	GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTTCCTCCT	2640
DB	2581	GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTTCCTCCT	2640
QY	2641	TACCTTCCCCCAGACTCGAATTTTTTCCCCCAACCCCAACTCTGAAGGTGAGAGTGAG	2700
DB	2641	TACCTTCCCCCAGACTCGAATTTTTTCCCCCAACCCCAACTCTGAAGGTGAGAGTGAG	2700

```

RESULT 3
US-09-744-314-14
; Sequence 14, Application US/09744314
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUGLER, Karl J.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 PCT
; CURRENT APPLICATION NUMBER: US/09/744,314
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/127,665
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq Ver. 3.0
; SEQ ID NO 14
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone No: 3768043
US-09-744-314-14

```

2701	QY	TTCCCCCTCAGGTTTGATTTCCACCTTCAAAATGAGCCCTGCTACTGAAACATCCAGAACCCACAG	2760
2701	Db	TTCCCCCTCAGGTTTGATTTCCACCTTCAAAATGAGCCCTGCTACTGAAACATCCAGAACCCACAG	2760
2761	QY	CAAGAAACCTGACAAATATATTTTGGCTCTCTTCAAAAGTAAATTTTGACTGATCTCATTTTCA	2820
2761	Db	CAAGAAACCTGACAAATATATTTTGGCTCTCTTCAAAAGTAAATTTTGACTGATCTCATTTTCA	2820
2821	QY	GTTTAAGTAACCTGCTGTACTTAAAGTGATACACTTTTGGCTCAAAATGGAAGCTTTAATGGA	2880
2821	Db	GTTTAAGTAACCTGCTGTACTTAAAGTGATACACTTTTGGCTCAAAATGGAAGCTTTAATGGA	2880
2881	QY	ATTATAAATTTCTCAGGATAGTATTTTCTGTAATTAAGATGATTTAAATATGAATCTTATGAG	2940
2881	Db	ATTATAAATTTCTCAGGATAGTATTTTCTGTAATTAAGATGATTTAAATATGAATCTTATGAG	2940
2941	QY	TAAATTTATTTTCAATTTTTATTTTAGACGGTATAACTATTTTCAATTTGATTAATCCACTATT	3000
2941	Db	TAAATTTATTTTCAATTTTTATTTTAGACGGTATAACTATTTTCAATTTGATTAATCCACTATT	3000
3001	QY	ATATAAACAAATAGTGGAGTTTTATATATGTAAATCTTTCAGGTGGGAGCCTTTAAATTC	3060
3001	Db	ATATAAACAAATAGTGGAGTTTTATATATGTAAATCTTTCAGGTGGGAGCCTTTAAATTC	3060
3061	QY	TGAAGTCTGTGCTTTATATGCAAGAACTGTATTTACTGTGTTTGTGCACAAATGTGAAAG	3120
3061	Db	TGAAGTCTGTGCTTTATATGCAAGAACTGTATTTACTGTGTTTGTGCACAAATGTGAAAG	3120
3121	QY	TAACTTTATGCTTTAAATAAATTAATGATTTGATTTAAAAAATAAAAAA	3164
3121	Db	TAACTTTATGCTTTAAATAAATTAATGATTTGATTTAAAAAATAAAAAA	3164

Db 61 CAAAATATGTGATGTGAACCTGCCGAGAACTTAGCGCGCGCGGCTCAGACACCGCTGC 120
QY 121 CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGCAACATCGGCGTG 180
Db 121 CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGCAACATCGGCGTG 180
QY 181 TGAGGCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
Db 181 TGAGGCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 241 CTCGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
Db 241 CTCGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
QY 301 GCTTGTGTGGCGAGTGTACTGCGGTGACCGCAGAGCAGCTCGAGCGCTATGGAGGAG 360
Db 301 GCTTGTGTGGCGAGTGTACTGCGGTGACCGCAGAGCAGCTCGAGCGCTATGGAGGAG 360
QY 361 CTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 361 CTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
QY 421 GCAGCACTACCTGAGGATGAGCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 480
Db 421 GCAGCACTACCTGAGGATGAGCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 480
QY 481 GTGGTATGTGCACTGTGATGAGATGATGAGATGATGAGATGATGAGATGATGAGATGAT 540
Db 481 GTGGTATGTGCACTGTGATGAGATGATGAGATGATGAGATGATGAGATGATGAGATGAT 540
QY 541 AGATCGGTAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
Db 541 AGATCGGTAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
QY 601 GGACTAAATGAAGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 660
Db 601 GGACTAAATGAAGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 660
QY 661 GAAGCGTATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
Db 661 GAAGCGTATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
QY 721 GAAGCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
Db 721 GAAGCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
QY 781 GATGAATCCTCTGTCTAGAAAAGACTTAAACAGAGAAATCTAAACATTTCTCAACAA 840
Db 781 GATGAATCCTCTGTCTAGAAAAGACTTAAACAGAGAAATCTAAACATTTCTCAACAA 840
QY 841 GATGAATCCTCTGTCTAGAAAAGACTTAAACAGAGAAATCTAAACATTTCTCAACAA 900
Db 841 GATGAATCCTCTGTCTAGAAAAGACTTAAACAGAGAAATCTAAACATTTCTCAACAA 900
QY 901 CTCAAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
QY 961 CGAGCGCTATAGCAATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
Db 961 CGAGCGCTATAGCAATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTGAAGTGA 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTGAAGTGA 1080
QY 1081 CAGAAATCAAAATGAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGA 1140
Db 1081 CAGAAATCAAAATGAGCTGAAGTATGGAAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGA 1140
QY 1141 AATCAATCAAGACCTGAGCTGGACACTTGCCTGCAATGATGAAGATCAGGCTCTGTGCTT 1200

Db 1141 AATCAATCAAGACCTGAGCTGGACACTTGCCTGCAATGATGAAGATCAGGCTCTGTGCTT 1200
QY 1201 GAAGAAGACACAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 GAAGAAGACACAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGAGCTTTGAAGAAATGATGATGATGATGATG 1320
Db 1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGAGCTTTGAAGAAATGATGATGATGATGATG 1320
QY 1321 TTAATGCTTCTTTTAAACCTTTAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TTAATGCTTCTTTTAAACCTTTAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 GAAGTTGATTAACAAAGAGAGAGCTTTACAGAGCATATTAATAATCTTCAAGACTCAACAA 1440
Db 1381 GAAGTTGATTAACAAAGAGAGAGCTTTACAGAGCATATTAATAATCTTCAAGACTCAACAA 1440
QY 1441 GCATCTTTGAGTCAAGAAACACATTTTGAAGATGAGATCAGAACTTCAAGAGATC 1500
Db 1441 GCATCTTTGAGTCAAGAAACACATTTTGAAGATGAGATCAGAACTTCAAGAGATC 1500
QY 1501 CTTAAAGTATGACTGAATTTATATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
Db 1501 CTTAAAGTATGACTGAATTTATATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
QY 1561 GTAGAGAAATTTATCGGTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATC 1620
Db 1561 GTAGAGAAATTTATCGGTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATC 1620
QY 1621 AGCATGCTCACTGAAGAGCTGGAGCTTATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 AGCATGCTCACTGAAGAGCTGGAGCTTATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 TTGAGAGAACTATTCATTTTATCAAGGAGAGATTTTCCCATGAGAGAGAGAGAGAGAG 1740
Db 1681 TTGAGAGAACTATTCATTTTATCAAGGAGAGATTTTCCCATGAGAGAGAGAGAGAGAG 1740
QY 1741 GATAATTTGGTGGAGCTGGAGTGTGAAGAACTCAATGATTTAAGGAGAGAGAGAGAG 1800
Db 1741 GATAATTTGGTGGAGCTGGAGTGTGAAGAACTCAATGATTTAAGGAGAGAGAGAGAG 1800
QY 1801 GCTCAACACAGACAAATTAATCTGAAACAGAGCTTTAAATTTGAACTTTTAGAGAGAG 1860
Db 1801 GCTCAACACAGACAAATTAATCTGAAACAGAGCTTTAAATTTGAACTTTTAGAGAGAG 1860
QY 1861 CTTATGCACTGATGTTCCAAATACAGCATTTGGCAGAGAGAGATTTCCCATATGCTCC 1920
Db 1861 CTTATGCACTGATGTTCCAAATACAGCATTTGGCAGAGAGAGATTTCCCATATGCTCC 1920
QY 1921 TCACCATTTGGTGGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCTCT 1980
Db 1921 TCACCATTTGGTGGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCT 1980
QY 1981 GAGGCTCACTGAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 1981 GAGGCTCACTGAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 2041 CCAGGGAGTCTCTGAGACCATCAGATTACCAATGAAGAGAGAGAGATCAAGCTGTGAT 2100
Db 2041 CCAGGGAGTCTCTGAGACCATCAGATTACCAATGAAGAGAGAGAGATCAAGCTGTGAT 2100
QY 2101 TTAACCGATCCTCATAGGCTCTCTCTGACACTGAGCTGTCTGTCTGTCTGTCTGTCTGTCT 2160
Db 2101 TTAACCGATCCTCATAGGCTCTCTCTGACACTGAGCTGTCTGTCTGTCTGTCTGTCTGTCT 2160
QY 2161 GACCGTAGGATGATGTTTCTCTCCGAGAGAGAGATCATATCTGATTTAGGCTTCTCTCTCA 2220
Db 2161 GACCGTAGGATGATGTTTCTCTCCGAGAGAGAGATCATATCTGATTTAGGCTTCTCTCTCA 2220
QY 2221 CAAAGGCAAGACAGATTTTCTTAAATTTCTGAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2280
Db 2221 CAAAGGCAAGACAGATTTTCTTAAATTTCTGAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2280

2281 AGTTTAAATGCTTCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2340
2281 AGTTTAAATGCTTCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2340
2341 AGTAGAAATGATACCAAGATGATCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2400
2341 AGTAGAAATGATACCAAGATGATCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2400
2401 GCTGAAATGATGAGGATGATGCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2460
2401 GCTGAAATGATGAGGATGATGCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 2460
2461 CCATTGTTCCAGTGGATGCAAGAGGCGCCATTCTTGAGAGAGAGGATCTCTTCCGCCCA 2520
2461 CCATTGTTCCAGTGGATGCAAGAGGCGCCATTCTTGAGAGAGAGGATCTCTTCCGCCCA 2520
2521 CTTCTCCAGAGGATGTTGGAGCTTCTGAGATTAATTTCCACAGAGGATTTCCCA 2580
2521 CTTCTCCAGAGGATGTTGGAGCTTCTGAGATTAATTTCCACAGAGGATTTCCCA 2580
2581 GGTCCACCACTGCTCCATTGCAATGAGAAATGCTTATCCACGAGGGGTTTCTCTCT 2640
2581 GGTCCACCACTGCTCCATTGCAATGAGAAATGCTTATCCACGAGGGGTTTCTCTCT 2640
2641 TACCTTCCCAAGACCTGGATTTTCCGCCACCCCAATCTGAGAGTGAAGTGA 2700
2641 TACCTTCCCAAGACCTGGATTTTCCGCCACCCCAATCTGAGAGTGAAGTGA 2700
2701 TTTCCCTCAGTTGATCTCCACCTTCAATGAGCTGCTACTGAAATCAATCCAGAACACAG 2760
2701 TTTCCCTCAGTTGATCTCCACCTTCAATGAGCTGCTACTGAAATCAATCCAGAACACAG 2760
2761 CAAGAAACCTGACAAATTTTGGCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
2761 CAAGAAACCTGACAAATTTTGGCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
2821 GTTTAGTAACTGCTGTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAACT 2880
2821 GTTTAGTAACTGCTGTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAACT 2880
2881 ATTATAATCTCAGATAGTATTTTGTAAATGATGATTTAAATGATGATTTAAATGATGAT 2940
2881 ATTATAATCTCAGATAGTATTTTGTAAATGATGATTTAAATGATGATTTAAATGATGAT 2940
2941 TAAATTAATTAATTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAA 3000
2941 TAAATTAATTAATTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAA 3000
3001 ATATAAATGATGAGGATTTTATATATGATGATTTTCAAGTGGGAGGCTTTAAATTC 3060
3001 ATATAAATGATGAGGATTTTATATATGATGATTTTCAAGTGGGAGGCTTTAAATTC 3060
3061 TGAAGTCTGCTTTATGCAAGAACTTAAATTTTAACTTAACTTAACTTAACTTAACTTAA 3120
3061 TGAAGTCTGCTTTATGCAAGAACTTAAATTTTAACTTAACTTAACTTAACTTAACTTAA 3120
3121 TAACTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3164
3121 TAACTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3164

RESULT 4

US-10-753-200-320
; Sequence 320, Application US/10753200
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/10753,200
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: 60/240,409

PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 320
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 3768043CB1
US-10-753-200-320

Query Match 100.0%; Score 3164; DB 55; Length 3164;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTTTAAATGCTTCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 60
DB 1 AGTTTAAATGCTTCTTTGGTAAATGATGGGTCAATGCTTCAGAAATGGATCC 60
QY 61 CAAATATGATGATGTAAGAACTGCGAGAACTAAGCGCGGCGGCTCAGACGCGTGC 120
DB 61 CAAATATGATGATGTAAGAACTGCGAGAACTAAGCGCGGCGGCTCAGACGCGTGC 120
QY 121 CTCAGATGTAAGTGTAAAGAGGCGGAGGAGGTGGTGGGGGCAACATGGGCGTG 180
DB 121 CTCAGATGTAAGTGTAAAGAGGCGGAGGAGGTGGTGGGGGCAACATGGGCGTG 180
QY 181 TGAGGCTGTGGGTGCGCGCTTCCAGCTTCCCGCGCGCGCTCCAGAGTGGTCCG 240
DB 181 TGAGGCTGTGGGTGCGCGCTTCCAGCTTCCCGCGCGCGCTCCAGAGTGGTCCG 240
QY 241 CTCGCTGTGGTGTCACTGCGCATTCGGGTTCAGACCCAGGCTGGGTGCTTCCACC 300
DB 241 CTCGCTGTGGTGTCACTGCGCATTCGGGTTCAGACCCAGGCTGGGTGCTTCCACC 300
QY 301 GCTTGTGTGGCGAGTGTACTGCGGTGACCGCGAGAGCAGCTCGAGCTATGGAGAG 360
DB 301 GCTTGTGTGGCGAGTGTACTGCGGTGACCGCGAGAGCAGCTCGAGCTATGGAGAG 360
QY 361 CTTGTGTACCCCTCAGCCCTACCTGGGTGCTGCTGGAGGAGCTACGAGAGTGTG 420
DB 361 CTTGTGTACCCCTCAGCCCTACCTGGGTGCTGCTGGAGGAGCTACGAGAGTGTG 420
QY 421 CGAGCACTACCTGAGAGTATGAGCCAGATCCTTATGTTTTCATCGGAATCG 480
DB 421 CGAGCACTACCTGAGAGTATGAGCCAGATCCTTATGTTTTCATCGGAATCG 480
QY 481 GTGGTATGTCAGCTGTTTATGGATTTTGTGTTCTCTTTTGTGGAGAGTTT 540
DB 481 GTGGTATGTCAGCTGTTTATGGATTTTGTGTTCTCTTTTGTGGAGAGTTT 540
QY 541 AGATCGGTAGAGTGGCTTTTACGTGGAGAGAGCAAACTTGGTGCACGCTTCT 600
DB 541 AGATCGGTAGAGTGGCTTTTACGTGGAGAGAGCAAACTTGGTGCACGCTTCT 600
QY 601 GGACTAAATGAGAAATGTAACCTACTTGAATAATTTAGCTTTTCAAAAGAGTAT 660
DB 601 GGACTAAATGAGAAATGTAACCTACTTGAATAATTTAGCTTTTCAAAAGAGTAT 660
QY 661 GAAGGCTATGAGTATGATCTTTAGAGATGCGAGCTTTGAGAGGAGGAGCAGAA 720
DB 661 GAAGGCTATGAGTATGATCTTTAGAGATGCGAGCTTTGAGAGGAGGAGCAGAA 720
QY 721 GAAGCAGCAAGTGTGGAGCAACCTGTGAAGAGCTGAACAGCTCCATTTGAACTT 780
DB 721 GAAGCAGCAAGTGTGGAGCAACCTGTGAAGAGCTGAACAGCTCCATTTGAACTT 780
QY 781 GATGAATCTCTGTCTAGAAAAAGACCTTAAACAGAGAAATCTTAAACATTTCTCA 840
DB 781 GATGAATCTCTGTCTAGAAAAAGACCTTAAACAGAGAAATCTTAAACATTTCTCA 840
QY 841 GATGAATCTCTGTCTAGAAAAAGATATACAGTCTCTAGAGATGATGAGTCAAAATCC 900

Db 841 GATGATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
QY 901 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTCCAAAGACATTTAAATCAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTCCAAAGACATTTAAATCAGTGAAGAA 960
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAAAGCCAT 1020
Db 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAAAGCCAT 1020
QY 1021 AAAAGCTTTTTCAGCAAGAGCTGAAGATGAGAAAGAGAGAGTGAAGTGAATCTTAATAAA 1080
Db 1021 AAAAGCTTTTTCAGCAAGAGCTGAAGATGAGAAAGAGAGAGTGAAGTGAATCTTAATAAA 1080
QY 1081 CAGAAAATAAATTTGAAGACTCCAAAGTACACGACAGAACAAAGTTCTCAATGATTAAGAA 1140
Db 1081 CAGAAAATAAATTTGAAGACTCCAAAGTACACGACAGAACAAAGTTCTCAATGATTAAGAA 1140
QY 1141 AATCATATCAAGACCTGACTGGACACTCTGCAATCTGCAATGATGAAGATCAGGCTGTGTCT 1200
Db 1141 AATCATATCAAGACCTGACTGGACACTCTGCAATCTGCAATGATGAAGATCAGGCTGTGTCT 1200
QY 1201 GAAGAGACACAAACGATGATGATACCTGGAAATAGAGTGAACAGTGAATCGGAAT 1260
Db 1201 GAAGAGACACAAACGATGATGATACCTGGAAATAGAGTGAACAGTGAATCGGAAT 1260
QY 1261 GGTGCTTACTTAGATAATCTCCAAAGAGAGCTTTCAAGAACTGATTCATGCTCTAAG 1320
Db 1261 GGTGCTTACTTAGATAATCTCCAAAGAGAGCTTTCAAGAACTGATTCATGCTCTAAG 1320
QY 1321 TTAATGCTCTTTAAACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TTAATGCTCTTTAAACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAATACTTCAGACTCAACAA 1440
Db 1381 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAATACTTCAGACTCAACAA 1440
QY 1441 GCATCTTCGAGTCAGAAACACACATTTTGAATGAGATCAGAGCTTCAGAGAA 1500
Db 1441 GCATCTTCGAGTCAGAAACACACATTTTGAATGAGATCAGAGCTTCAGAGAA 1500
QY 1501 CTTAAAGTAATGACTGAAATATATCAAGAAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
Db 1501 CTTAAAGTAATGACTGAAATATATCAAGAAAATGAAATGAAATGAAATGAAATGAA 1560
QY 1561 GTAGAGAAAATATTCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAGATC 1620
Db 1561 GTAGAGAAAATATTCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAGATC 1620
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCAAGATCTTGAAGAGAA 1680
Db 1621 AGCCATGCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCAAGATCTTGAAGAGAA 1680
QY 1681 TTGGAGAGAACTATTCATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAGACAT 1740
Db 1681 TTGGAGAGAACTATTCATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAGACAT 1740
QY 1741 GATTAATGTTGGAGCTGGAGTCTGAAAGAGAACTCAATGATTTAAGGAAAGAAAT 1800
Db 1741 GATTAATGTTGGAGCTGGAGTCTGAAAGAGAACTCAATGATTTAAGGAAAGAAAT 1800
QY 1801 GCTCAACAGACAAAATTAACAGTGAACAGAGCTTAAATTTGAATTTGAAGAAAAGAT 1860
Db 1801 GCTCAACAGACAAAATTAACAGTGAACAGAGCTTAAATTTGAATTTGAAGAAAAGAT 1860
QY 1861 CCTATGCTCTGATGCTCAATATCAGATTTGGCAGAGAGATTTCCCATATGCTCC 1920
Db 1861 CCTATGCTCTGATGCTCAATATCAGATTTGGCAGAGAGATTTCCCATATGCTCC 1920
QY 1921 TCACCATTTGGGTTGGCTTCATCTGAACAAAGAGCTTTTCTCTCTCTCCCAACTTTGTTG 1980

Db 1921 TCACCATTTGGGTTGGCTTCATCTGAACAAAGAGCTTTTCTCTCTCTCCCAACTTTGTTG 1980
QY 1981 GAGGTCCTCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGAGAGCTTCAGAGGC 2040
Db 1981 GAGGTCCTCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGAGAGCTTCAGAGGC 2040
QY 2041 CCAGGGAATCCTCTGACCATCAGATTTACCAATGAAAGAGGAGAAATCAAGCTGTGATAG 2100
Db 2041 CCAGGGAATCCTCTGACCATCAGATTTACCAATGAAAGAGGAGAAATCAAGCTGTGATAG 2100
QY 2101 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGGGACCAAG 2160
Db 2101 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGGGACCAAG 2160
QY 2161 GACCGTAGGATGATGTTTCTCGCCAGGACCAATCATATCTCTGATTCAGCCCTTCCTCCA 2220
Db 2161 GACCGTAGGATGATGTTTCTCGCCAGGACCAATCATATCTCTGATTCAGCCCTTCCTCCA 2220
QY 2221 CAAAGGCAAGACAGATTTTGTCTTAATCTGTAGACTCTGTGGACCAAGCAATCTCAG 2280
Db 2221 CAAAGGCAAGACAGATTTTGTCTTAATCTGTAGACTCTGTGGACCAAGCAATCTCAG 2280
QY 2281 AGTTTAAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAAATGGAATCC 2340
Db 2281 AGTTTAAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAAATGGAATCC 2340
QY 2341 AGTAGAAAATGATACCAAGATGATCTTGGTAATTTAAATGTGCTGATTCATCTCTCCCT 2400
Db 2341 AGTAGAAAATGATACCAAGATGATCTTGGTAATTTAAATGTGCTGATTCATCTCTCCCT 2400
QY 2401 GCTGAAAATCAAGCCACTGGCCCTGCTTGTTCCTCCACCTCTTCTCAATCAGAGGT 2460
Db 2401 GCTGAAAATGAGGCCACTGGCCCTGCTTGTTCCTCCACCTCTTCTCAATCAGAGGT 2460
QY 2461 CCAATGTTTCCAGTGAATGCAAGAGGCCAATCTTGAGAGAGAGACCTCTCTTTCCCTCCCA 2520
Db 2461 CCAATGTTTCCAGTGAATGCAAGAGGCCAATCTTGAGAGAGAGACCTCTCTTTCCCTCCCA 2520
QY 2521 CCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCCAAGAGGATTTCCCA 2580
Db 2521 CCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCCAAGAGGATTTCCCA 2580
QY 2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCT 2640
Db 2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCT 2640
QY 2641 TACCTTCCCAAGAGCTGGATTTTCCCTCCACCCCACTTCGAGTGAAGTGAAGTGAAG 2700
Db 2641 TACCTTCCCAAGAGCTGGATTTTCCCTCCACCCCACTTCGAGTGAAGTGAAGTGAAG 2700
QY 2701 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAACATCCAGAACCAAG 2760
Db 2701 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAACATCCAGAACCAAG 2760
QY 2761 CAAAGAACCTGCAATATTTTCTCTTCAAAAGTAAATTTGATCTGATCTCATTTTCA 2820
Db 2761 CAAAGAACCTGCAATATTTTCTCTTCAAAAGTAAATTTGATCTGATCTCATTTTCA 2820
QY 2821 GTTTAAAGTAACTGCTGTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAA 2880
Db 2821 GTTTAAAGTAACTGCTGTTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAA 2880
QY 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATGAAGTATTTAAATGAATCTTATGAG 2940
Db 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATGAAGTATTTAAATGAATCTTATGAG 2940
QY 2941 TAAATTAATTTCAATTTTATTTAGAGGATATACTATTTCAATTTGATTAATCCACTATT 3000
Db 2941 TAAATTAATTTCAATTTTATTTAGAGGATATACTATTTCAATTTGATTAATCCACTATT 3000
QY 3001 ATATAAACAATAGTGGAGTTTATATATGATTTCTTCAAGTGGGAGGCTTTAAATTC 3060
Db 3001 ATATAAACAATAGTGGAGTTTATATATGATTTCTTCAAGTGGGAGGCTTTAAATTC 3060

Query March	100.0%	Score 3164	DB 81	Length 3164	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 3164	Conservative	0	Mismatches 0	Indels 0	Gaps 0
Qy	1	AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCCACCAAAAATAAAGCTT	60		
Db	1	AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCCACCAAAAATAAAGCTT	60		
Qy	61	CAAAATATGTGATGTGAAAACTGCCAGAACTAAGCGGGCCGGGCTCAGACACGGCTGC	120		
Db	61	CAAAATATGTGATGTGAAAACTGCCAGAACTAAGCGGGCCGGGCTCAGACACGGCTGC	120		
Qy	121	CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGGAGGTGGTGGGGGCAAAATCGGGCGTG	180		
Db	121	CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGGAGGTGGTGGGGGCAAAATCGGGCGTG	180		
Qy	181	TGAGGCGTGTGGTGTCCCGGTTTCCCAAGTCTCCCGCCAGCCGCGCTCCAAGTGTCTCG	240		
Db	181	TGAGGCGTGTGGTGTCCCGGTTTCCCAAGTCTCCCGCCAGCCGCGCTCCAAGTGTCTCG	240		
Qy	241	CTCCGGTTGGTTGTCAAGTGGGCAATTCGGGTTCCAGACCCCAAGGTCGGTGTCTCCACC	300		
Db	241	CTCCGGTTGGTTGTCAAGTGGGCAATTCGGGTTCCAGACCCCAAGGTCGGTGTCTCCACC	300		
Qy	301	GCTTGTGTGGCCAGTGTACTGCGGTGACCCCAAGAGAGCGCTCGACGCTATGGAGAG	360		
Db	301	GCTTGTGTGGCCAGTGTACTGCGGTGACCCCAAGAGAGCGCTCGACGCTATGGAGAG	360		
Qy	361	CTGTGTGTACCCCTCAGCCCTACCTGCGGCTGGTCTCTGGAGGAGCTACGCAGAGTTCTG	420		
Db	361	CTGTGTGTACCCCTCAGCCCTACCTGCGGCTGGTCTCTGGAGGAGCTACGCAGAGTTCTG	420		
Qy	421	GCAGCACTACCTGAGAGTATGAGACAGATGAGAAATCCTTATGTTTTCCATCGGAACTG	480		
Db	421	GCAGCACTACCTGAGAGTATGAGACAGATGAGAAATCCTTATGTTTTCCATCGGAACTG	480		
Qy	481	GTGGTATGTGCAGCTGTATTGGATTTTTTGTGTCTCCCTTTTTTCTGGAGAGTATT	540		
Db	481	GTGGTATGTGCAGCTGTATTGGATTTTTTGTGTCTCCCTTTTTTCTGGAGAGTATT	540		
Qy	541	AGATCGGTTAGGAGTCGGCTTTTACGTGGGAAGAGAGCAAAAACCTTGGTGCACACGTTTCT	600		

1681 TTGGAGAACTATTTCATCTTATCAGGGCAGATTATTTCCCATGAGAAAAGACACAT 1740
1681 TTGGAGAACTATTTCATCTTATCAGGGCAGATTATTTCCCATGAGAAAAGACACAT 1740
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
1801 GCTCAACAGACAAAATTTAACTCAACAGAGCTTAAATTTGAACCTTTTGAAGAAAAGAT 1860
1801 GCTCAACAGACAAAATTTAACTCAACAGAGCTTAAATTTGAACCTTTTGAAGAAAAGAT 1860
1861 CTTATGCACTCGATGTTTCCAAATACAGCAATTTGCACAGAGCAATCCCATATGTTGCC 1920
1861 CTTATGCACTCGATGTTTCCAAATACAGCAATTTGCACAGAGCAATCCCATATGTTGCC 1920
1921 TCACCAATTTGGTGGCTTCATCTGAACAGAGCTTTTCTCTCTCCCACTTTGTTG 1980
1921 TCACCAATTTGGTGGCTTCATCTGAACAGAGCTTTTCTCTCTCCCACTTTGTTG 1980
1981 GAGGTTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
1981 GAGGTTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
2041 CCAGGAATCCTCTGACCATCAGATTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
2041 CCAGGAATCCTCTGACCATCAGATTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
2101 TTAAACCGATCCTCATAGGGCTCCCTCTGACACTGGGCTCTCTCTCACTCCATGGACACG 2160
2101 TTAAACCGATCCTCATAGGGCTCCCTCTGACACTGGGCTCTCTCTCACTCCATGGACACG 2160
2161 GACCGTAGATGATGTTCTCGCCAGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
2161 GACCGTAGATGATGTTCTCGCCAGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
2221 CAAAGGCAACAGAGATTTTGTCTAATCTGTGAGACTGCTGGACCAAGCAACTCAGA 2280
2221 CAAAGGCAACAGAGATTTTGTCTAATCTGTGAGACTGCTGGACCAAGCAACTCAGA 2280
2281 AGTTTAAATAGCTCTTTTGGATTAATAGGATGGGTCAATGCTTCAGAAATGGAATCC 2340
2281 AGTTTAAATAGCTCTTTTGGATTAATAGGATGGGTCAATGCTTCAGAAATGGAATCC 2340
2341 AGTAGAAATCATACCAAGATGATCTTGTAAATTTAAATGTGCTGATTCATCTCTCCCT 2400
2341 AGTAGAAATCATACCAAGATGATCTTGTAAATTTAAATGTGCTGATTCATCTCTCCCT 2400
2401 GCTGAAATCAAGCACTGCTGCTGCTTTGTTGCTTCTCCCACTTGTCTCAATCAGAGT 2460
2401 GCTGAAATCAAGCACTGCTGCTGCTTTGTTGCTTCTCCCACTTGTCTCAATCAGAGT 2460
2461 CCATTGTTTCCAGTGGATCAAGAGGCCCAATCTTGAAGAGAGAGCTCTCTTCCCCCA 2520
2461 CCATTGTTTCCAGTGGATCAAGAGGCCCAATCTTGAAGAGAGAGCTCTCTTCCCCCA 2520
2521 CCTCCTCCAGAGCCATGTTGGAGCTTCTCGAGATTAATTTTCAACAAAGGATTTCCCA 2580
2521 CCTCCTCCAGAGCCATGTTGGAGCTTCTCGAGATTAATTTTCAACAAAGGATTTCCCA 2580
2581 GGTCCACCACTGCTCCATTTGCAATCAGAAATGCTCTATCCAGAGGGGTTTTCTCTCT 2640
2581 GGTCCACCACTGCTCCATTTGCAATCAGAAATGCTCTATCCAGAGGGGTTTTCTCTCT 2640
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAACCCCACTTCTGAAAGTAGAAGTAG 2700
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAACCCCACTTCTGAAAGTAGAAGTAG 2700
2701 TTCCCTCAGGTTTGAATCCACTCAATGAGCCTCTCTGAACTCCAGACACAG 2760
2701 TTCCCTCAGGTTTGAATCCACTCAATGAGCCTCTCTGAACTCCAGACACAG 2760

2761 CAAGAAACCTGACAAATATTTTGTCTCTCTCAAAAGTAATTTTGACTGATCTCATTTTCA 2820
2761 CAAGAAACCTGACAAATATTTTGTCTCTCTCAAAAGTAATTTTGACTGATCTCATTTTCA 2820
2821 GTTTAAGTAACCTGCTGTTACTTAAGTGATTAACACTTTTGTCTCAAAATGGAAGCTTAATGA 2880
2821 GTTTAAGTAACCTGCTGTTACTTAAGTGATTAACACTTTTGTCTCAAAATGGAAGCTTAATGA 2880
2881 ATTATTAATCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGAACTCTTAATGAG 2940
2881 ATTATTAATCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGAACTCTTAATGAG 2940
2941 TAAATTAATCTCAATTTTATTTAGACGGTATAACTATTTCAATTTGATTAAATCCACTATT 3000
2941 TAAATTAATCTCAATTTTATTTAGACGGTATAACTATTTCAATTTGATTAAATCCACTATT 3000
3001 ATATAACAATAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTTTAAATTC 3060
3001 ATATAACAATAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTTTAAATTC 3060
3061 TGAAGTCTGCTCTTTATGCCAAGAACTGATTTACTGTTGTTGTCGACAAATGTGAAAG 3120
3061 TGAAGTCTGCTCTTTATGCCAAGAACTGATTTACTGTTGTTGTCGACAAATGTGAAAG 3120
3121 TAACTTTATGCTTTAAATAAATTAATGATTTGATTTTAAAAAATAAAAA 3164
3121 TAACTTTATGCTTTAAATAAATTAATGATTTGATTTTAAAAAATAAAAA 3164

RESULT 6

US-60-324-185-1645
; Sequence 1645, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lai, Preeti
; APPLICANT: Deep, Diph
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 1645
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 039946.18
US-60-324-185-1645

Query Match 99.8%; Score 3159.2; DB 89; Length 3164;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3161; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGTTTAAATCCATGAGAGACGCAATTTTAAAGTGATTTACCAAAAAATAAAGCTT 60
DB 1 AGGTTTAAATCCATGAGAGACGCAATTTTAAAGTGATTTACCAAAAAATAAAGCTT 60
QY 61 CAAAAATATGATGTGAAACTGCCAGAACTAAGGCGGGCGGGCTCAGACAGGCTGC 120
DB 61 CAAAAATATGATGTGAAACTGCCAGAACTAAGGCGGGCGGGCTCAGACAGGCTGC 120
QY 121 CTCAGAGTGAAGTGTAAACAGAGGCGGAGGAGGTTGGGGGACAAATGGGCTG 180
DB 121 CTCAGAGTGAAGTGTAAACAGAGGCGGAGGAGGTTGGGGGACAAATGGGCTG 180
QY 181 TGAGGCGCTGTGGGTGCCCGGCTTCCCAAGCTCCCCCGCAGCCCGCTCCACAGTGGTCCG 240
DB 181 TGAGGCGCTGTGGGTGCCCGGCTTCCCAAGCTCCCCCGCAGCCCGCTCCACAGTGGTCCG 240

241 QY CTCGGTGTGTTGTCAGCTGGCATTCCGGTTCCAGACCCAGGCTCGGTGTTCTCCACC 300
241 Db CTCGGTGTGTTGTCAGCTGGCATTCCGGTTCCAGACCCAGGCTCGGTGTTCTCCACC 300
301 QY GCTTGTGTGGCAGTGTACTCGGGTGACCGGCGAGAGAGCCTCGACGCTATGGAGGAG 360
301 Db GCTTGTGTGGCAGTGTACTCGGGTGACCGGCGAGAGAGCCTCGACGCTATGGAGGAG 360
361 QY CTTGGTCTACCCCTCAGCCCTACCTGGGCTGGTCTCGGAGGAGCTACGAGAGTTGTG 420
361 Db CTTGGTCTACCCCTCAGCCCTACCTGGGCTGGTCTCGGAGGAGCTACGAGAGTTGTG 420
421 QY GCAGCACTACTGAGAGTATGAGACCAAGATGAGAACTCTTATGGTTTGCATCGGAACCTG 480
421 Db GCAGCACTACTGAGAGTATGAGACCAAGATGAGAACTCTTATGGTTTGCATCGGAACCTG 480
481 QY GTGGTATGTCAGCTGTATTGGATTTTTTGTGTTCTCCCTTTTTTGTGGAGAGTTT 540
481 Db GTGGTATGTCAGCTGTATTGGATTTTTTGTGTTCTCCCTTTTTTGTGGAGAGTTT 540
541 QY AGATCGGTTAGGAGTGGCTTTACGTGGGAGAGAGCAAAAATCTGGTGCACGCTTTCT 600
541 Db AGATCGGTTAGGAGTGGCTTTACGTGGGAGAGAGCAAAAATCTGGTGCACGCTTTCT 600
601 QY GGACTAATCGAAGAAATCTAACTACTTGAATAATTTAGCCCTTATCAAAAAGAGTAT 660
601 Db GGACTAATCGAAGAAATCTAACTACTTGAATAATTTAGCCCTTATCAAAAAGAGTAT 660
661 QY GAAGGCTATGAAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAA 720
661 Db GAAGGCTATGAAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAA 720
721 QY GAAGCAGAGTTGGAGGCAACTGTGAAGAGCTGAACAGGTCCTAATCTGAGCTTGAG 780
721 Db GAAGCAGAGTTGGAGGCAACTGTGAAGAGCTGAACAGGTCCTAATCTGAGCTTGAG 780
781 QY GATGAATCCTCTGTCTAGAAAAGAGCTTAAACACAGAGAAATCTAAACATCTCAACAA 840
781 Db GATGAATCCTCTGTCTAGAAAAGAGCTTAAACACAGAGAAATCTAAACATCTCAACAA 840
841 QY GATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
841 Db GATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
901 QY CTCAAATCACAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATAGAGTGAAGAA 960
901 Db CTCAAATCACAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATAGAGTGAAGAA 960
961 QY CGAGCGCTATAGCAATAAAGATGCTTTGAATGAATAATCTCAATCTCAGACAAGCCAT 1020
961 Db CGAGCGCTATAGCAATAAAGATGCTTTGAATGAATAATCTCAATCTCAGACAAGCCAT 1020
1021 QY AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
1021 Db AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
1081 QY CAGAAAATAACATTTGAAGCTCCAAAGTGCAGCAGAAACAAGTCTGTAATGATAAAGAA 1140
1081 Db CAGAAAATAACATTTGAAGCTCCAAAGTGCAGCAGAAACAAGTCTGTAATGATAAAGAA 1140
1141 QY AATCAGATCAGACCTGACTGGACATCTGCCAATGATGAAGATCAGGCTGCTGCTT 1200
1141 Db AATCAGATCAGACCTGACTGGACATCTGCCAATGATGAAGATCAGGCTGCTGCTT 1200
1201 QY GAAGAAGACACAACCGGATGATGAATCTGGAATTTAGAGTGAACAGTGAATCGGAAT 1260
1201 Db GAAGAAGACACAACCGGATGATGAATCTGGAATTTAGAGTGAACAGTGAATCGGAAT 1260
1261 QY GGTGCTTACTTAGATAATCTCCAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
1261 Db GGTGCTTACTTAGATAATCTCCAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
1321 QY TTAATATGCTTTCTTTAAAAACCTTAGAAGGAGAGAAACCAAAATTTATATTCTGCTCT 1380

1321 Db TTAATATGCTTTCTTTAAAAACCTTAGAAGGAGAGAAACCAAAATTTATATTCTGCTCT 1380
1381 QY GAAGTTGATAAAAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
1381 Db GAAGTTGATAAAAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
1441 QY GCATCTTTGCACTCAGAAAAACACACATTTTGAATAAGATCAGAAAGCTTCAACAGAAA 1500
1441 Db GCATCTTTGCACTCAGAAAAACACACATTTTGAATAAGATCAGAAAGCTTCAACAGAAA 1500
1501 QY CTTAAAGTAATGACTGAATATATCAAGAAATGAATGAATCTCCAGAGAAATTAACA 1560
1501 Db CTTAAAGTAATGACTGAATATATCAAGAAATGAATGAATCTCCAGAGAAATTAACA 1560
1561 QY GTAGAGGAAAATTTACGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAGATC 1620
1561 Db GTAGAGGAAAATTTACGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAGATC 1620
1621 QY AGCCATGCACTGAGAGCTCGAGACCTATAGAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
1621 Db AGCCATGCACTGAGAGCTCGAGACCTATAGAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
1681 QY TTGGAGAGAACTATTCTTATCAAGGCGAGATTTATTTCCCATGAGAAAAAGCACAT 1740
1681 Db TTGGAGAGAACTATTCTTATCAAGGCGAGATTTATTTCCCATGAGAAAAAGCACAT 1740
1741 QY GATAATTCGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAAT 1800
1741 Db GATAATTCGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAAT 1800
1801 QY GCTCAACAACAGCAAAAATTAACCTGAAACAGAGCTTTAAATTTGAACCTTTAGAAAAACAT 1860
1801 Db GCTCAACAACAGCAAAAATTAACCTGAAACAGAGCTTTAAATTTGAACCTTTAGAAAAACAT 1860
1861 QY CTTATGCACTCGATGTTCCAAATAACAGATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
1861 Db CTTATGCACTCGATGTTCCAAATAACAGATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
1921 QY TCACCATTTGGGTTGGCCTTCATCTGAAACAGAGCTTTCTCTCTCTCCTCAACTTTGTTG 1980
1921 Db TCACCATTTGGGTTGGCCTTCATCTGAAACAGAGCTTTCTCTCTCTCCTCAACTTTGTTG 1980
1981 QY GAGGCTCCACTCAGACTCTCACCTTTGCTTCAGGGGAGGAGGAGGCTCAGAGGC 2040
1981 Db GAGGCTCCACTCAGACTCTCACCTTTGCTTCAGGGGAGGAGGAGGCTCAGAGGC 2040
2041 QY CCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
2041 Db CCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
2101 QY TTAACCGATCCTCATAGGCTCCCTCTGACACTGGGTCTGTCTCACCTCCATGGAGCCAG 2160
2101 Db TTAACCGATCCTCATAGGCTCCCTCTGACACTGGGTCTGTCTCACCTCCATGGAGCCAG 2160
2161 QY GACCGTAGATGATGTTCTTCGCGCAGAGCAATCATATCTGATTCAGCCCTTCTCCA 2220
2161 Db GACCGTAGATGATGTTCTTCGCGCAGAGCAATCATATCTGATTCAGCCCTTCTCCA 2220
2221 QY CAAAGGCAAGACAGATTTTGTCTTAATTTCTGTAGACTGTCTGGACCAAGAGAACTCAGA 2280
2221 Db CAAAGGCAAGACAGATTTTGTCTTAATTTCTGTAGACTGTCTGGACCAAGAGAACTCAGA 2280
2281 QY AGTTTAAATATGCTTTTGGATTAATTAATGAGGTCAATGCTTTCAGAAATGAATTC 2340
2281 Db AGTTTAAATATGCTTTTGGATTAATTAATGAGGTCAATGCTTTCAGAAATGAATTC 2340
2341 QY AGTAGAAATGATACCAAGAGATGATCTTGTGTAATTTAAATGTCCTGATTCATCTCCT 2400
2341 Db AGTAGAAATGATACCAAGAGATGATCTTGTGTAATTTAAATGTCCTGATTCATCTCCT 2400
2401 QY GCTGAAATGAAGCCACTGGCCCTGGCTTTGTTCTCCACCTTTGCTCCCAATCAGAGT 2460

1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGAAGTAAATAA 1080
1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGAAGTAAATAA 1080
1081 CAGAAATTAACATTTGAAGACTCCAAAGTACACGAGAGCAAGTCTTGAATGATAAGAA 1140
1081 CAGAAATTAACATTTGAAGACTCCAAAGTACACGAGAGCAAGTCTTGAATGATAAGAA 1140
1141 AATCAGATCAAGACCTGACTGAGACCTTGCCAAATCATGAAGATCAGGCTGTGCTT 1200
1141 AATCAGATCAAGACCTGACTGAGACCTTGCCAAATCATGAAGATCAGGCTGTGCTT 1200
1201 GAAGAGACACAAACGGATGATATAACCTGGAAATAGAAAGTGAACAGTGAATCGGAAAT 1260
1201 GAAGAGACACAAACGGATGATATAACCTGGAAATAGAAAGTGAACAGTGAATCGGAAAT 1260
1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGCTTTGAAGAACTGATTCATGCTGTAAG 1320
1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGCTTTGAAGAACTGATTCATGCTGTAAG 1320
1321 TTAATGCTTCTTTTAAACCTTTAGAGAGAGAAAGAAACCAATTTATATTCAGTTGCT 1380
1321 TTAATGCTTCTTTTAAACCTTTAGAGAGAGAAAGAAACCAATTTATATTCAGTTGCT 1380
1381 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTAATAAT-CTTCAGACTCAACA 1439
1381 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACA 1440
1440 AGCATCTTTGCAGTCAGAAACACACATTTTGAANAATGAGAATCAGAAGCTTCAACAGAA 1499
1441 AGCATCTTTGCAGTCAGAAACACACATTTTGAANAATGAGAATCAGAAGCTTCAACAGAA 1500
1500 ACTTAAAGTAACTGAATATATCAAGAAATGAANAATGAACCTCCACAGAAATTAAC 1559
1501 ACTTAAAGTAACTGAATATATCAAGAAATGAANAATGAACCTCCACAGAAATTAAC 1560
1560 AGTAGAGAAATATCGGTAGAGAGAGAGAACTTTCTAAAGTAGATGAAGAAGAT 1619
1561 AGTAGAGAAATATCGGTAGAGAGAGAGAACTTTCTAAAGTAGATGAAGAAGAT 1620
1620 CAGCCATGCCACTGAAGAGCTGGAGACTTATAGAAAGCGAGCCAAAGATCTTGAAGAGA 1679
1621 CAGCCATGCCACTGAAGAGCTGGAGACTTATAGAAAGCGAGCCAAAGATCTTGAAGAGA 1680
1680 ATTGGAGAGAACTATTCATCTTATCAAGGCGAGATTTCCCATCAGAGAAAGAGCACA 1739
1681 ATTGGAGAGAACTATTCATCTTATCAAGGCGAGATTTCCCATCAGAGAAAGAGCACA 1740
1740 TGATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAA 1800
1800 TGCTCACAACAGACAAATAATTAAGTGAACAGAGCTTAAATTTGAACTTTAGAAAGA 1859
1801 TGCTCACAACAGACAAATAATTAAGTGAACAGAGCTTAAATTTGAACTTTAGAAAGA 1860
1860 TCCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTCC 1919
1861 TCCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTCC 1920
1920 CTCACCAATGGGTGGCCTTCATCTGAACAGAGCTTTTCTCTCTCTCCATCTTGT 1979
1921 CTCACCAATGGGTGGCCTTCATCTGAACAGAGCTTTTCTCTCTCTCCATCTTGT 1980
1980 GGAGGTCACACTCAGACTCTCACCTTTGCTTCCAGGGGGAGGAGAGAGGCTCACGAGG 2039
1981 GGAGGTCACACTCAGACTCTCACCTTTGCTTCCAGGGGGAGGAGAGAGGCTCACGAGG 2040
2040 CCCAGGGAATCTCTGGACCATCAGATTAACCAAGAGAGAGAAATCAAGCTGTGATAG 2099
2041 CCCAGGGAATCTCTGGACCATCAGATTAACCAAGAGAGAGAAATCAAGCTGTGATAG 2100
2100 GTTAACCGATCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATGGGACCA 2159

RESULT 8

2101 GTTAACCGATCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGGGACCA 2160
2160 GGACGCTAGAGATGATGTTTCTCCGCGCAGACCAATCATATCTCGATTGAGCCCTTCCTCC 2219
2161 GGACGCTAGAGATGATGTTTCTCCGCGCAGACCAATCATATCTCGATTGAGCCCTTCCTCC 2220
2220 ACAAAGCACAAGACAGATTTGTTCTAATTTCTGTAGACTGTCTGGACCAAGCAACTCAG 2279
2221 ACAAAGCACAAGACAGATTTGTTCTAATTTCTGTAGACTGTCTGGACCAAGCAACTCAG 2280
2280 AAGTTTAAATATGCTCTCTTTGGATAAATGAGTGGGTCAATGCTTCAGAAAATGGAAATC 2339
2281 AAGTTTAAATATGCTCTCTTTGGATAAATGAGTGGGTCAATGCTTCAGAAAATGGAAATC 2340
2340 CAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCC 2399
2341 CAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCC 2400
2400 TGCTGAAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTGCTCCAAATCAGAGG 2459
2401 TGCTGAAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTGCTCCAAATCAGAGG 2460
2460 TCCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGACCTCTCTTCCCTCC 2519
2461 TCCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGACCTCTCTTCCCTCC 2520
2520 ACCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTTCCCAAGAGGATTTCCCTCC 2579
2521 ACCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTTCCCAAGAGGATTTCCCTCC 2580
2580 AGGTCCACACCTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTTCTCTCC 2639
2581 AGGTCCACACCTGCTCCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTTCTCTCC 2640
2640 TTACCTTCCCAAGACCTGGATTTTCCCAAGAGGCTTCTACTGAAATCAGAACCCACA 2699
2641 TTACCTTCCCAAGACCTGGATTTTCCCAAGAGGCTTCTACTGAAATCAGAACCCACA 2700
2700 GTTCCCTCAGGTTGATTTCCACCTTCAAATGAGGCTGCTACTGAAATCAGAACCCACA 2759
2701 GTTCCCTCAGGTTGATTTCCACCTTCAAATGAGGCTGCTACTGAAATCAGAACCCACA 2760
2760 GCAAGAAACCTGACAAATATTTTGTCTCTCTCAAAGTAAATTTTGACTGATCTCAATTTT 2819
2761 GCAAGAAACCTGACAAATATTTTGTCTCTCTCAAAGTAAATTTTGACTGATCTCAATTTT 2820
2820 AGTTTAAAGTAACTGCTGTTTAAAGTAACTGATTAACATTTTGTCTCAAATTTGAAGCTTAA 2879
2821 AGTTTAAAGTAACTGCTGTTTAAAGTAACTGATTAACATTTTGTCTCAAATTTGAAGCTTAA 2880
2880 AATTATAATTTCTCAGGATAGTATTTTGTAAATPAAAGATGATTTAAATATGAATCTTATGA 2939
2881 AATTATAATTTCTCAGGATAGTATTTTGTAAATPAAAGATGATTTAAATATGAATCTTATGA 2940
2940 GTAATATTTTCAAATTTTATTTAGACGATTAACATTTTCAAATTTGATTAATCCACTAT 2999
2941 GTAATATTTTCAAATTTTATTTAGACGATTAACATTTTCAAATTTGATTAATCCACTAT 3000
3000 TATATAAACAATAGTGGAGTTTTATATATATATATATATATATATATATATATATATATAT 3059
3001 TATATAAACAATAGTGGAGTTTTATATATATATATATATATATATATATATATATATATAT 3060
3060 CTGAAGTCTGTCTTTATGCAAGAACTGATTTACTGTGTTTGTGGCAATGTAATGTA 3119
3061 CTGAAGTCTGTCTTTATGCAAGAACTGATTTACTGTGTTTGTGGCAATGTAATGTA 3120
3120 GTAACCTTATGCTTAAATAAATATAGTTGATTTTAAAAA 3164
3121 GTAACCTTATGCTTAAATAAATATAGTTGATTTTAAAAA 3165

PCT-US02-25465-51

; Sequence 51, Application PC/TUS0225465

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: BARROSO, Ines

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: BECHA, Shanya D.

; APPLICANT: BOROWSKI, Mark L.

; APPLICANT: DUGGAN, Brendan M.

; APPLICANT: ELLIOTT, Vicki S.

; APPLICANT: EMERLING, Brooke M.

; APPLICANT: FORSYTHE, Ian J.

; APPLICANT: GIETZEN, Kimberly J.

; APPLICANT: GORVAD, Ann E.

; APPLICANT: GRIFFIN, Jennifer C.

; APPLICANT: GURURAJAN, Rajagopal

; APPLICANT: HAFALIA, April J.A.

; APPLICANT: ISON, Craig H.

; APPLICANT: KABLE, Amy E.

; APPLICANT: KHAN, Farrah A.

; APPLICANT: LEE, Sally

; APPLICANT: LEE, Soo Yeun

; APPLICANT: LI, Joana X.

; APPLICANT: REDDY, Roopa

; APPLICANT: RICHARDSON, Thomas W.

; APPLICANT: SPRAGUE, William W.

; APPLICANT: SWARNAKAR, Anita

; APPLICANT: TANG, Y. Tom

; APPLICANT: WARREN, Bridget A.

; APPLICANT: XU, Yuming

; APPLICANT: YAO, Monique G.

; APPLICANT: YUE, Henry

; APPLICANT: YUE, Huibin

; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH

; FILE REFERENCE: PP-1126 PCT

; CURRENT APPLICATION NUMBER: PCT/US02/25465

; CURRENT FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: US 60/311,017

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/313,070

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/313,071

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/314,678

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/316,692

; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/317,913

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/322,182

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/340,747

; PRIOR FILING DATE: 2001-12-07

; PRIOR APPLICATION NUMBER: US 60/342,761

; PRIOR FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: US 60/369,129

; PRIOR FILING DATE: 2002-03-29

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PERL Program

; SEQ ID NO 51

; LENGTH: 3108

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 7506167CBI

PCT-US02-25465-51

Query Match

Best Local Similarity 91.0%; Score 2880.4;

Matches 3023; Conservative 0; Mismatches 6; Indels 129; Gaps 1;

DB 1; Length 3108;

Pred. No. 0;

Matches 3023; Conservative 0; Mismatches 6; Indels 129; Gaps 1;

QY	1	AGGTTTATCCATGAAGAAGACAGCAATTTTAAAAAGTGTTATTCACCAAAAAATAAGCTT	60
Db	72	AGGTTTATCCATGAAGAAGACAGCAATTTTAAAAAGTGTTATTCACCAAAAAATAAGCTT	131
QY	61	CAAAATATGTGATGTGAAAACTGCCAGAACTAAAGCGGGCCGGGCTCAGACAGCGCTGC	120
Db	132	CAAAATATGTGATGTGAAAACTAAAGCGGGCCGGGCTCAGACAGCGCTGC	191
QY	121	CTCAGGATGTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGACAAATAGGCGCTG	180
Db	192	CTCAGGATGTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGACAAATAGGCGCTG	251
QY	181	TGAGGCGCTGTGGTGCCCGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG	240
Db	252	TGAGGCGCTGTGGTGCCCGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG	311
QY	241	CTCCGGTGGTGTACAGTGCGCATTTCCGGTTCAGACCCAAAGGCTCGTGTCTCCACC	300
Db	312	CTCCGGTGGTGTACAGTGCGCATTTCCGGTTCAGACCCAAAGGCTCGTGTCTCCACC	371
QY	301	GCTTGTGTGGCCAGTGTACTGCGGTGACCGCAGAGCAGCTCGACCTATGGAGGAG	360
Db	372	GCTTGTGTGGCCAGTGTACTGCGGTGACCGCAGAGCAGCTCGACCTATGGAGGAG	431
QY	361	CCTGGTGCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGACAGTTGTG	420
Db	432	CCTGGTGCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGACAGTTGTG	491
QY	421	GCAGCACTACCTGAGAGTATGAGACCAGATGAGATCCCTTATGTTTCCATCGGAATG	480
Db	492	GCAGCACTACCTGAGAGTATGAGACCAGATGAGATCCCTTATGTTTCCATCGGAATG	551
QY	481	GTGCTATGTGCAGCTGTTATTTGGATTTTGTGTTTCTCTTTTGTGGAGAACTTT	540
Db	552	GTGCTATGTGCAGCTGTTATTTGGATTTTGTGTTTCTCTTTTGTGGAGAACTTT	611
QY	541	AGATCGGTAGGAGTGGCTTTTACGTCGGAGAGACCAAACTTGGTCCAGCGTTTCT	600
Db	612	AGATCGGTAGGAGTGGCTTTTACGTCGGAGAGACCAAACTTGGTCCAGCGTTTCT	671
QY	601	GGACTAATTTGAAGAAAAATGTAACCTACTTTGAAAAATTTAGCCTTATTCAAAAAGAT	660
Db	672	GGACTAATTTGAAGAAAAATGTAACCTACTTTGAAAAATTTAGCCTTATTCAAAAAGAT	731
QY	661	GAAGGCTATGAAGTAGATCATCTTTAGAGATGCCAGCTTTGAGAGAGGCGCAGAGAA	720
Db	732	GAAGGCTATGAAGTAGATCATCTTTAGAGATGCCAGCTTTGAGAGAGGCGCAGAGAA	791
QY	721	GAAGCAGCAAGTTTGGAGGCACTCTGAAAAAGCTGAAACAGGTCCCAATTTCTGAACCTGAG	780
Db	792	GAAGCAGCAAGTTTGGAGGCACTCTGAAAAAGCTGAAACAGGTCCCAATTTCTGAACCTGAG	851
QY	781	GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACCTTCTCAACAA	840
Db	852	GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACCTTCTCAACAA	911
QY	841	GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACCTTCTCAACAA	900
Db	912	GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACCTTCTCAACAA	971
QY	901	CTCAATTCACAAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA	960
Db	972	CTCAATTCACAAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA	1031
QY	961	CGACGGCTATAGCAATAAAGATGCTTTCGAATGAAATTTCTCAACTTCAGACAAAGCAT	1020
Db	1032	CGACGGCTATAGCAATAAAGATGCTTTCGAATGAAATTTCTCAACTTCAGACAAAGCAT	1091
QY	1021	AAACAGCTTTTTCAGCAAGAGCTGAAATGATGAAAGGAGAGAGTGAAGTAAATAAA	1080
Db	1092	AAACAGCTTTTTCAGCAAGAGCTGAAATGATGAAAGGAGAGAGTGAAGTAAATAAA	1151

1081 CAGAAAATAACATTTGAGACTCCAAAGTACACGCAACAAAGTTCTGATGATRAAGAA 1140
1152 CAGAAAATAACATTTGAGACTCCAAAGTACACGCAACAAAGTTCTGATGATRAAGAA 1211
1141 AATCAGATCAAGACCCCTGACTGACACTCTCCAAATGATGAAGATCAGGCTGTGCTT 1200
1212 AATCAGATCAAGACCCCTGACTGACACTCTCCAAATGATGAAGATCAGGCTGTGCTT 1271
1201 GAAGAAGACACAAAGGATGATTAACCTCGAATTAAGATGAACGATGAATCGGAAT 1260
1272 GAAGAAGACACAAAGGATGATTAACCTCGAATTAAGATGAACGATGAATCGGAAT 1331
1261 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
1332 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1391
1321 TTAATGCTTCTTTAAAGACCTTAGAGGAGAAAGAAACCAATTTATATTCAGTTGCT 1380
1392 TTAATGCTTCTTTAAAGACCTTAGAGGAGAAAGAAACCAATTTATATTCAGTTGCT 1451
1381 GAAGTTGATTAAGAAAGGAGCTTACAGAGCATATTAAGAACTTTCAGACTCAACAA 1440
1452 GAAGTTGATTAAGAAAGGAGCTTACAGAGCATATTAAGAACTTTCAGACTCAACAA 1511
1441 GCATCTTCAGATCAGAAACACACATTTTGAAGATGAAGATCAGAGCTTCAACAGAAA 1500
1512 GCATCTTCAGATCAGAAACACACATTTTGAAGATGAAGATCAGAGCTTCAACAGAAA 1571
1501 CTTAAAGTAACTGATGAATTAATCAAGAAATGAAGATGAAGATCAGAGCTTCAACAG 1560
1572 CTTAAAGTAACTGATGAATTAATCAAGAAATGAAGATGAAGATCAGAGCTTCAACAG 1631
1561 GTAGAGAAATTAATCGGTTAGAGAAAGAGAAACTTTCTAAAGTAGATGAAGAAATC 1620
1632 GTAGAGAAATTAATCGGTTAGAGAAAGAGAAACTTTCTAAAGTAGATGAAGAAATC 1691
1621 AGCCATCCACTGAAGAGCTGAGACCTATAGAAAGGAGCCAAAGATCTTGAAGAGAA 1680
1692 AGCCATCCACTGAAGAGCTGAGACCTATAGAAAGGAGCCAAAGATCTTGAAGAGAA 1751
1681 TTGAGAGAACTATTCATTTCTTAAGGCGCAGATTAATTTCCCATGAGAAAAAGCACAT 1740
1752 TTGAGAGAACTATTCATTTCTTAAGGCGCAGATTAATTTCCCATGAGAAAAAGCACAT 1811
1741 GATAATTTGGTGGCAGCTCGGAATCTGGAAGAACTCAATGATTTAGGAAAGAAAT 1800
1812 GATAATTTGGTGGCAGCTCGGAATCTGGAAGAACTCAATGATTTAGGAAAGAAAT 1871
1801 GCTCACAACACAGAAAAATTAAGTGAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 1860
1872 GCTCACAACACAGAAAAATTAAGTGAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 1931
1861 CCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCCTC 1920
1932 CCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTGGCAGAGAGAG 1967
1921 TCACCATTTGGGTTGGCTTCTATCTGAACCAAGAGCTTTCTCTCTCTCCCACTTTGTTG 1980
1968 ----- 1967
1981 GAGGTCCTCAGACTCTCACCCTTTCAGGGGGGAGAGAGAGCTCAGAGGC 2040
1968 -----AGAGGCTCAGAGGC 1982
2041 CCAGGAACTCTTGACCTCAGATTAACCAATGAAGAGAGAACTCAAGCTGTGATAGG 2100
1983 CCAGGAACTCTTGACCTCAGATTAACCAATGAAGAGAGAACTCAAGCTGTGATAGG 2042
2101 TTAACCGATCTCATAGGCTCTCTGACATGGGCTCTGTCACCTCCATCGGACCGAG 2160
2043 TTAACCGATCTCATAGGCTCTCTGACATGGGCTCTGTCACCTCCATCGGACCGAG 2102
2161 GACCGTAGGATGATTTCTCCCGCAGGACAAATCATATCTGATTCAGGCCCTCTCCCA 2220

2103 GACCGTAGGATGATTTCTCCCGCAGGACAAATCATATCTGATTCAGGCCCTCTCTCCA 2162
2221 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCAAGAGACTCAGA 2280
2163 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCAAGAGACTCAGA 2222
2281 AGTTTAAATATGCTCTTTCTGGATAAAATGATGGGTCAATGCCCTTCAGAAATGAATCC 2340
2223 AGTTTAAATATGCTCTTTCTGGATAAAATGATGGGTCAATGCCCTTCAGAAATGAATCC 2282
2341 AGTAAATATGATACAAAGATGATCTTTGGTAAATTTAAATGTGCCGTATCATCTCCCT 2400
2283 AGTAAATATGATACAAAGATGATCTTTGGTAAATTTAAATGTGCCGTATCATCTCCCT 2342
2401 GCTGAAATGAAGCCACCTGGCCCTTGTCTCCACCTTTGTCTCCAAATCAGAGGT 2460
2343 GCTGAAATGAAGCCACCTGGCCCTTGTCTCCACCTTTGTCTCCAAATCAGAGGT 2402
2461 CCATTTGTTTCCAGTGGATGCAAGAGGCCCAATTTGAGAGAGAGGACCTCTTTCCCCCA 2520
2403 CCATTTGTTTCCAGTGGATGCAAGAGGCCCAATTTGAGAGAGAGGACCTCTTTCCCCCA 2462
2521 COTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTAATTTCCACCAAGAGGATTTCCCA 2580
2463 COTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTAATTTCCACCAAGAGGATTTCCCA 2522
2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGTCTATCCACGAGGGGTTTCTCTCT 2640
2523 GGTCCACACCTGCTCCATTTGCAATGAGAAATGTCTATCCACGAGGGGTTTCTCTCT 2582
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGAGGATTTCTGAGGTAGAGTAG 2700
2583 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGAGGATTTCTGAGGTAGAGTAG 2642
2701 TTCCCTCAGGTTTGAATTTCCACCTTCAATGAGCTGCTACTGAACTCCAGAACCCAG 2760
2643 TTCCCTCAGGTTTGAATTTCCACCTTCAATGAGCTGCTACTGAACTCCAGAACCCAG 2702
2761 CAAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTTGAATCTCATTTTCA 2820
2703 CAAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTTGAATCTCATTTTCA 2762
2821 GTTTAAGTAACTGCTGTTTACTTAAGTATACACTTTTGTCTCAAAATGAAGCTTAATCGA 2880
2763 GTTTAAGTAACTGCTGTTTACTTAAGTATACACTTTTGTCTCAAAATGAAGCTTAATCGA 2822
2881 ATTATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTAAATATGAATCTTATGAG 2940
2823 ATTATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTAAATATGAATCTTATGAG 2882
2941 TAAATTTTCAATTTTATTTTAGCGGTATACTATTTCAATTTGATTAATCCACTATT 3000
2893 TAAATTTTCAATTTTATTTTAGCGGTATACTATTTCAATTTGATTAATCCACTATT 2942
3001 ATATAAACAATAGTGGAGTTTATATATGTAATTTCTTTCAGGTGGGAGGCTTTAAATTC 3060
2943 ATATAAACAATAGTGGAGTTTATATATGTAATTTCTTTCAGGTGGGAGGCTTTAAATTC 3002
3061 TGAAGTCTGCTCTTTATGCGCAAGAACTGATTTTACTGTTGTTGTCGCAAAATCTGGAAG 3120
3003 TGAAGTCTGCTCTTTATGCGCAAGAACTGATTTTACTGTTGTTGTCGCAAAATCTGGAAG 3062
3121 TAACTTTTATGCTTAAATAAATTAATGATTTGATTTAAAGA 3158
3063 TAACTTTTATGCTTAAATAAATTAATGATTTGATTTAAAGA 3100

RESULT 9

US-60-500-315-11859/c
; Sequence 11859, Application US/60500315
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING HUMAN G-PROTEIN COUPLED RECEPTOR PROTEINS, METHODS
 ; FILE REFERENCE: CL001484
 ; CURRENT APPLICATION NUMBER: US/60/500,315
 ; NUMBER OF SEQ ID NOS: 69978
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11859
 ; LENGTH: 122644
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(122644)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-60-500-315-11859

Query Match 89.6%; Score 2836.4; DB 107; Length 122644;
 Best Local Similarity 94.0%; P-red. No. 0;
 Matches 2975; Conservative 25; Mismatches 151; Indels 13; Gaps 5;
 QY 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 60
 Db 80139 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 80080
 QY 61 CAAAATATGTGATGTAAGAACTGCCAGAACTAAGCGCGCGGCTCAGACAGCGCTGC 120
 Db 80079 CAAAATATGTGATGTAAGAACTGCCAGAACTAAGCGCGCGGCTCAGACAGCGCTGC 80020
 QY 121 CTCAGGATTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGGACAAATGGGCGCTG 180
 Db 80019 CTCAGGATTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGGACAAATGGGCGCTG 79960
 QY 181 TGAGGCGCTGTGGTGGCCCGCTTCCCGAGCTCCCGCGAGCGCGCTCCACAGTGTGCTCG 240
 Db 79959 TGAGGCGCTGTGGTGGCCCGCTTCCCGAGCTCCCGCGAGCGCGCTCCACAGTGTGCTCG 79900
 QY 241 CTCGGGTGTGGTGTACAGTGGGATTCGGGTTCAGACCCAAAGGTGGTGTTCACAC 300
 Db 79899 CTCGGGTGTGGTGTACAGTGGGATTCGGGTTCAGACCCAAAGGTGGTGTTCACAC 79840
 QY 301 GCTTGTGTGGCAGTGTACTCGGTGACCGCAGAGCAGCTCGAGCTATGAGGAG 360
 Db 79839 GCTTGTGTGGCAGTGTACTCGGTGACCGCAGAGCAGCTCGAGCTATGAGGAG 79780
 QY 361 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTGAGGAGCTACGAGAGTGTG 420
 Db 79779 CTTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTGAGGAGCTACGAGAGTGTG 79720
 QY 421 GCAGCACTACCTGAGAGTATGACCCAGATGAGATTCCTTATGTTTCCATCGGAACTG 480
 Db 79719 GCAGCACTACCTGAGAGTATGACCCAGATGAGATTCCTTATGTTTCCATCGGAACTG 79660
 QY 481 GTGGTATGTGCAGCTGTATTGGATTTTGTGTTTCTCCTTTTGTGGAGAGTGT 540
 Db 79659 GTGGTATGTGCAGCTGTATTGGATTTTGTGTTTCTCCTTTTGTGGAGAGTGT 79600
 QY 541 AGATCGTTAGGAGTGGCTTTACGTGGGAAGAGCAGCAAACTTGGTGCACGCTTCT 600
 Db 79599 AGATCGTTAGGAGTGGCTTTACGTGGGAAGAGCAGCAAACTTGGTGCACGCTTCT 79540
 QY 601 GGACTAATTCAGAAAATGTAACTTACCTTGTAAATTTAGCCTTATTCAAAAGAGTAT 660
 Db 79539 GGACTAATTCAGAAAATGTAACTTACCTTGTAAATTTAGCCTTATTCAAAAGAGTAT 79480
 QY 661 GAAGGCTATGAGTAGAGTCACTTTAGAGGATGCCAGCTTTAGAGAGGAGCAGAGAA 720
 Db 79479 GAAGGCTATGAGTAGAGTCACTTTAGAGGATGCCAGCTTTAGAGAGGAGCAGAGAA 79420
 QY 721 GAAGCACGAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGTCCAAATCTGAACTTGG 780
 Db 79419 GAAGCACGAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGTCCAAATCTGAACTTGG 79360

QY 781 GATGAATCTCTGTCTAGAAAAAGACTTAAACAGAGAAATCTAAACATCTCAACAA 840
 Db 79359 GATGAATCTCTGTCTAGAAAAAGACTTAAACAGAGAAATCTAAACATCTCAACAA 79300
 QY 841 GATGAATCTGATGCGGATATTTCAAAGATATACAGCTCTCTAGAGATGAGTCAAAATCC 900
 Db 79299 GATGAATCTGATGCGGATATTTCAAAGATATACAGCTCTCTAGAGATGAGTCAAAATCC 79240
 QY 901 CTCAATTCACAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 960
 Db 79239 CTCAATTCACAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 79180
 QY 961 CGAGCGCTATAGCAATAAAGATGCTTTCAATGAAAAATTTCAACTTCAGACAAGCCAT 1020
 Db 79179 YGAGCGCTATAGCAATAAAGATGCTTTCAATGAAAAATTTCAACTTCAGACAAGCCAT 79120
 QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGTAGTGAATTAATAA 1080
 Db 79119 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGTAGTGAATTAATAA 79060
 QY 1081 CAGAAATACATTTGAAGACTCCAAAGTACACGCAAGAACAGTTCTGATGATAAGAA 1140
 Db 79059 CAGAAATACATTTGAAGACTCCAAAGTACACGCAAGAACAGTTCTGATGATAAGAA 79000
 QY 1141 AATCAGATCAAGACCCCTGACTGACACTTGCCTAATGATGAAGATCAGGCTGTGCTT 1200
 Db 78999 AATCAGATCAAGACCCCTGACTGGACACTTGCCTAATGATGAAGATCAGGCTGTGCTT 78940
 QY 1201 GAGAGACACACAGGATGATGATTAAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1260
 Db 78939 GAGAGACACACAGGATGATGATTAAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 78880
 QY 1261 GGTGCTTACTTAGATTAATCTCCAAAGGAGCTTTGAAGAAACTGATTCATGCTGCTAG 1320
 Db 78879 GGTGCTTACTTAGATTAATCTCCAAAGGAGCTTTGAAGAAACTGATTCATGCTGCTAG 78820
 QY 1321 TTAATGCTTTCTTAAACCTTTAGAGGAGAGAGAAACCAATTTATTTATTCAGTTGCT 1380
 Db 78819 TTAATGCTTTCTTAAACCTTTAGAGGAGAGAGAAACCAATTTATTTATTCAGTTGCT 78760
 QY 1381 GAAGTTGATAAAACAAAGGAGAGCTTTACAGACATATTAAAAATCTTCAGACTCAACAA 1440
 Db 78759 GAAGTTGATAAAACAAAGGAGAGCTTTACAGACATATTAAAAATCTTCAGACTCAACAA 78700
 QY 1441 GCATCTTTGAGTCAAGAAACACACATTTTGAAGATGAGATCAGAGCTTCAACAGAA 1500
 Db 78699 GCATCTTTGAGTCAAGAAACACACATTTTGAAGATGAGATCAGAGCTTCAACAGAA 78640
 QY 1501 CTTAAAGTAACTGACTGATTAATCAAGAAATGAATGAATGAATGAATGAATGAATGAATGA 1560
 Db 78639 CTTAAAGTAACTGACTGATTAATCAAGAAATGAATGAATGAATGAATGAATGAATGAATGA 78580
 QY 1561 GTAGAGAAATTTATCGTTTAGAGAAAGAGAAACTTTCTTAAAGTATGATGAAGATC 1620
 Db 78579 GTAGAGAAATTTATCGTTTAGAGAAAGAGAAACTTTCTTAAAGTATGATGAAGATC 78520
 QY 1621 AGCCATCCCTCAGAGCTGGAGACTATAGAAAGCAGGCAAGATCTTGAAGAGAGAA 1680
 Db 78519 AGCCATCCCTCAGAGCTGGAGACTATAGAAAGCAGGCAAGATCTTGAAGAGAGAA 78460
 QY 1681 TTGAGAGAACTATTCATTTTATCAAGGGCAGATTTTCCCATGAGAAAAAGACAT 1740
 Db 78459 TTGAGAGAACTATTCATTTTATCAAGGGCAGATTTTCCCATGAGAAAAAGACAT 78400
 QY 1741 GATTAATTTGGTGGCAGCTCGGAATGCTGAAGAACTCAATGATTTAAGGAAGAAAT 1800
 Db 78399 GATTAATTTGGTGGCAGCTCGGAATGCTGAAGAACTCAATGATTTAAGGAAGAAAT 78340
 QY 1801 GCTCACACAGACAAATAATTAACCTGAACAGAGCTTTAAATTTGAACTTTTAGAAAAAGAT 1860
 Db 78339 GCTCACACAGACAAATAATTAACCTGAACAGAGCTTTGAAATTTGAACTTTTAGAAAAAGAT 78280

1861 CTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCAATCCCCATGATGTC 1920
Db CTTAATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCAATCCCCATGATGTC 78220
1921 TCACCAATGGGTTGGGCTTCATCTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Db TCACCAATGGGTTGGGCTTCATCTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 78160
1981 GAGGGTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
Db GAGGATCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGAGGCTCAGAGGC 78100
2041 CAGAGGAATCCTCTGACACATCAGATACCAATGAAAGAGGAGATCAAGCTGTGATAGG 2100
Db CAGAGGAATCCTCTGACACATCAGATACCAATGAAAGAGGAGATCAAGCTGTGATAGG 78040
2101 TTAACCGATCCTCATAGGGCTCCTCTGACACTGGGCTCTGTGACCTCCATGGGACAG 2160
Db TTAATCGATCCTCATAGGGCTCCTCTGACACTGGGCTCCTGTGACCTCCGCTGACAG 77980
2161 GACCGTAGGATGATGTTCTCCGCGCAGGACATCATATCTGATTCAGCCCTTCTCTCA 2220
Db GACCGTAGGATGATGTTCTCCGCGCAGGACATCATATCTGATTCAGCTTCTCTCA 77920
2221 CAAAGGCAAGACAGATTTGTTCTTAATCTGTPAGACTCTGAGCAGCAGACATCAGA 2280
Db CAAAGGCAAGACAGATTTGTTCTTAATCTGTPAGACTCTGAGCAGCAGACATCAGA 77860
2281 AGTTTAAATGCTTCTTTGGATAAAATGGATGGGCTCAATGCTTTCAGAAATGGAATCC 2340
Db AGTTTAAATGCTTCTTTGGATAAAATGGATGGGCTCAATGCTTTCAGAAATGGAATCC 77800
2341 AGTAGAAATGATACCAAGATGATCTGTTAAATGTTAAATGCTGATTCATCTCTCCCT 2400
Db AGTAGAAATGATACCAAGATGATCTGTTAAATGTTAAATGCTGATTCATCTCTCCCT 77740
2401 GCTGAAATGAAACCACTGGCCCTGCTTGTTCCTCCACTCTTCTCTCAATCAGAGT 2460
Db GCTGAAATGAAACCACTGGCCCTGCTTGTTCCTCCACTCTTCTCTCAATCAGAGT 77730
2461 CAAATGTTCCAGTGATGCAAGAGGCCATCTTTCAGAGAGGAGCTCTCTTTCCTCCCA 2520
Db CAAATGTTCCAGTGATGCAAGAGGCCATCTTTCAGAGAGGAGCTCTCTTTCCTCCCA 77620
2521 CTTCTCCAGAGCCATGTTTGGAGTCTCTGAGATTAATTTCCACCAAGAGGATTTCCCA 2580
Db CTTCTCCAGAGCCATGTTTGGAGTCTCTGAGATTAATTTCCACCAAGAGGATTTCCCA 77610
2581 GGTCCACACCTGCTCCTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCT 2640
Db GGTCCACACCTGCTCCTGCAATGAGAAATGCTATCCACCGAGGGGTTTACCTCT 77500
2641 TACCTTCCCAAGACCTGGATTTTCCCTCCCACTTCTGAGTGAAGTGAAGTGAAG 2700
Db TACCTTCCCAAGACCTGGATTTTCCCTCCCACTTCTGAGTGAAGTGAAGTGAAG 77442
2701 TTTCCCTCAGTTGATTCACCTCAATGAGCCTGCTACTGAAATCAGAACACACAG 2760
Db TTTCCCTCAGTTGATTCACCTCAATGAGCCTGCTACTGAAATCAGAACACACAG 77382
2761 CAAAGAACCTGCAATATTTTCTCTCTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
Db CAAAGAACCTGCAATATTTTCTCTCTCAAAAGTAAATTTGACTGATCTCATTTTCA 77322
2821 GTTTAAAGTAACTGCTGTTACTTAAGTAAATTTGCTTCAAAATGAACTTAATGGA 2880
Db GTTTAAAGTAACTGCTGTTACTTAAGTAAATTTGCTTCAAAATGAACTTAATGGA 77263
2881 ATTATATCTCAGGATGATTTTGAATTAAGATGATTTAAATGAACTTAATGGA 2940
Db ATTATATCTCAGGATGATTTTGAATTAAGATGATTTAAATGAACTTAATGGA 77203
2941 TAAATTAATTTCAATTT-----TATTTAGAGCGGTATAACTATTTCAT--TTGATTAATC 2993

77202 TAAATCATTTCCATTTTATTTATTTAGATCATATACTTTTAACTTTGGTGAATATC 77143
Qy CACTATTATATAACATAGTGGAGTTTATATATGTAATCTTTAGGTGGAGGCTT 3053
Db CACTTTTAGAGAAACAATAGTGGAGTTTATATATGTAATCTTTAGGTGGAGGCTT 77083
3054 TAAATTTCTGAAGTCT--GTGTCTTTATGCAAGAACTGTATTTACTGTGTGTGACA 3110
Db TAAATTTCTGAAGTCTTTATGCTTTAAATAAATTTAGTTGATTTA 3154
Qy AATGTGAAGTAACTTTATGCTTTAAATAAATTTAGTTGATTTA 77023
Db AATGTGAAGTAACTTTATGCTTTAAATAAATTTAGTTGATTTA 76979

RESULT 10
US-09-948-941-107
; Sequence 107, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107
; LENGTH: 3659
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-107

Query Match 87.9%; Score 2780.2; DB 39; Length 3659;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2931; Conservative 0; Mismatches 118; Indels 30; Gaps 5;

Qy 90 CTAAGGGGGGGGGCTCAGACAGCCCTGCTCAGATGTAAAGTGTAAACAGAGGCC 149
Db 57 CTTGTGGGGTGGGGCTCGGACCTGCTCGGATGTAAAGTGTAAACAGAGGGTC 116
Qy 150 AGGGAGAGTGGTGGGGGACAAACATGGGCTGTGAGGCTGTGGGTGCGGCTTCCCGAG 209
Db 117 GGGATGGG-----CAGCTAGGCTGTGAGGCTGTGGGTGCGGCTTCCCGAG 166
Qy 210 CTCCCCCGCAGCCGCTCCACAGTGTGCTCGGTGGTGTGCTGCTGCTGCTGCTGCTGCT 269
Db 167 CTCCCCCGCAGCCGCTCCACAGTGTGCTCGGTGGTGTGCTGCTGCTGCTGCTGCTGCT 222
Qy 270 GTTCCAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
Db 223 GTTCCAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 282
Qy 320 ACTCGGTGACCGCCAGAGCAGCTCGACGTATGAGAGGCTGCTGCTGCTGCTGCTGCTGCT 379
Db 283 ACTGCGGACCCAGAGAGAGCTTGGCGCTATGAGAGGCTGCTGCTGCTGCTGCTGCTGCT 342
Qy 380 CTAAGTGGGCTGCTGCTGAGGAGCTACGAGAGTGTGCGAGCTACCTGAGAGTA 439
Db 343 CGTATTTGGGCTGCTGCTGAGGAGCTACGAGAGTGTGCGAGCTACCTGAGAGTA 402
Qy 440 TGAGACAGATGAGAACTCTTATGTTTTCATCGAACTGGTGTGATGTGCTGCTGCTGCT 499
Db 403 TGAGACAGATGAGAACTCTTATGTTTTCATCGAACTGGTGTGATGTGCTGCTGCTGCT 462
Qy 500 TTGATTTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559
Db 463 TTGATTTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522
Qy 560 TTTAGTGGGAGAGAGCAAAAATTGCTGCAAGCTTTTCTGCTGCTGCTGCTGCTGCTGCT 619

Db 523 TTTATGTGGGACGAGAGAAAGCTTCTTAATGCTTTCGACTAATTGAGAAATA 582
Qy 620 GTAACTACTTGAATAATTTAGCTTATTCATAAGAGATGAGAGCTATGAGTAGT 679
Db 583 GTAACTACTTGAATAATTTAGCTTATTCATAAGAGATGAGAGCTATGAGTAGT 642
Qy 680 CATCTTTAGAGTCCAGCTTTGAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAG 739
Db 643 CATCTTTAGAGTCCAGCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 699
Qy 740 CAACCTGTGAAGAGTGAACAGCTGCAATCTGAACTTGAAGATGAATTCCTGTCTAG 799
Db 700 CAACCTGTGAAGAGTGAACAGCTGCAATCTGAACTTGAAGATGAATTCCTGTCTAG 759
Qy 800 AAAAAAGCTTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
Db 760 AAAAAAGCTTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
Qy 860 TTTCAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCAAAATAGCTG 919
Db 820 TTTCAAAAGGATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCAAAATAGCTG 879
Qy 920 AAGCCAAATCATCTGCAAGACATTTAAATCAGTGAAGAGAGAGAGAGAGAGAG 979
Db 880 AAGCCAAATGACCTTCAAGATATTTCAATGATGAAGAGAGAGAGAGAGAGAG 939
Qy 980 AAGTGTCTTGAATGAATAATCTCACTTCAGACAGAGAGAGAGAGAGAGAGAG 1039
Db 940 AAGTGTCTTGAATGAATAATCTCACTTCAGAGAGAGAGAGAGAGAGAGAGAG 996
Qy 1040 AAGCTGAAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1099
Db 997 AAGCTGAAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
Qy 1100 ACTCAAAGTACAG 1159
Db 1057 ACTCAAAGTACAG 1116
Qy 1160 CTGGACATCTGCAATGATGAAGATCAGCTGCTGCTGAGAGAGAGAGAGAGAG 1219
Db 1117 CTGAAGCTTTGTAAGATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
Qy 1220 ATGATAACCTGGAATTTAGAGTGAACAGTGAATCGGAGAGAGAGAGAGAGAGAG 1279
Db 1177 ATGATAACCTGGAATTTAGAGTGAACAGTGAATCGGAGAGAGAGAGAGAGAGAG 1236
Qy 1280 CTCCAAAGAGCTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339
Db 1237 CTCCAAAGAGAGCTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
Qy 1340 CCTTAGAAG 1399
Db 1297 CCTTAGAAG 1356
Qy 1400 AAGAGCTTACAG 1459
Db 1357 AAGAGCTTACAG 1416
Qy 1460 ACACACATTTGAAATGAGATCAGAGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAG 1519
Db 1417 ACACACATTTGAAATGAGATCAGAGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAG 1476
Qy 1520 TATATCAAG 1579
Db 1477 TATATCAAG 1536
Qy 1580 TAG 1639
Db 1537 TAG 1596
Qy 1640 TGAG 1699

Db 1597 TGGAGACCTATAGAAAGCGAGCCAAAGATCTTCAAGAGAAATTTGGAGAGAACTATTTCATT 1656
Qy 1700 CTTATCAAGGGAGATATTTCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1759
Db 1657 CTTATCAAGGGAGATATTTCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1716
Qy 1760 GGAATGCTGAAG 1819
Db 1717 GGAATGCTGAAG 1776
Qy 1820 TAACTGAAG 1879
Db 1777 TAACTGAAG 1836
Qy 1880 CAAATACAGATTTGGCAG 1939
Db 1837 CAAATACAGATTTGGCAG 1896
Qy 1940 CATCTGAAG 1999
Db 1897 CATCTGAAG 1956
Qy 2000 CACCTTTGCTTCCAGGGGAG 2059
Db 1957 CACCTTTGCTTCCAGGGGAG 2016
Qy 2060 ATCAGATTTACCAATGAAG 2119
Db 2017 ATCAGATTTACCAATGAAG 2076
Qy 2120 CTCCCTCTGACAGCTGGTCTCTGCTCACTTCCATGGGAGAGAGAGAGAGAGAGAGAGAGAG 2179
Db 2077 CTCCCTCTGACAGCTGGTCTCTGCTCACTTCCATGGGAGAGAGAGAGAGAGAGAGAGAGAG 2136
Qy 2180 CTCGGCAG 2239
Db 2137 CTCGGCAG 2196
Qy 2240 GTTCTAATCTGCTAG 2299
Db 2197 GTTCTAATCTGCTAG 2256
Qy 2300 TGGATAAATGAG 2359
Db 2257 TGGATAAATGAG 2316
Qy 2360 ATGATCTTGTGTAATTTAAATGCTGATCATCTCTCCCTGCTGAGAGAGAGAGAGAGAGAGAG 2419
Db 2317 ATGATCTTGTGTAATTTAAATGCTGATCATCTCTCCCTGCTGAGAGAGAGAGAGAGAGAGAG 2376
Qy 2420 GCCTTGGCTTTGCTTCCCTCAGCTTCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2479
Db 2377 GCCTTGGCTTTGCTTCCCTCAGCTTCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2436
Qy 2480 CAG 2539
Db 2437 CAG 2496
Qy 2540 TGGAG 2599
Db 2497 TGGAG 2556
Qy 2600 TGGAG 2659
Db 2557 TGGAG 2616
Qy 2660 GATTTTTCCTCCAG 2719
Db 2617 GATTTTTCCTCCAG 2676
Qy 2720 CACCTTCAAAATGAG 2779
Db 2677 CACCTTCAAAATGAG 2736

1520 TATATCAAGAAATGAATGAATGAACTCCACAGAAATTTAACTAGAGAGAAATTTATCGGT 1579
1541 TATATCAAGAAATGAATGAATGAACTCCACAGAAATTTAACTAGAGAGAAATTTATCGGT 1600
1580 TAGAGAAAGAGAGAAATCTTTCTAAAGTAGATGAAGAGATCAGCCATCCCACTGAAGAGC 1639
1601 TAGAGAAAGAGAGAAATCTTTCTAAAGTAGATGAAGAGATCAGCCATCCCACTGAAGAGC 1660
1640 TGGAGACTATAG 1699
1661 TGGAGACTATAG 1720
1700 CTTATCAAGGGCAGATTTATTTCCCATGAGAAAAAGCAGATGATTAATTTGGTTGGCAGCTC 1759
1721 CTTATCAAGGGCAGATTTATTTCCCATGAGAAAAAGCAGATGATTAATTTGGTTGGCAGCTC 1780
1760 GGAATGCTGAAG 1819
1781 GGAATGCTGAAG 1840
1820 TAACTGAAG 1879
1841 TAACTGAAG 1900
1880 CAAATACAG 1939
1901 CAAATACAG 1960
1940 CATCTGAAG 1999
1961 CATCTGAAG 2020
2000 CACCTTTGCTCCAGGGGAG 2059
2021 CACCTTTGCTCCAGGGGAG 2080
2060 ATCAGATTTACCAATGAAG 2119
2081 ATCAGATTTACCAATGAAG 2140
2120 CTCCTCTGACACTGGGTCTCTGTCACCTCCATGGGACAGGACCGTAGGATGATGTTTC 2179
2141 CTCCTCTGACACTGGGTCTCTGTCACCTCCATGGGACAGGACCGTAGGATGATGTTTC 2200
2180 CTCGGCCAG 2239
2201 CTCGGCCAG 2260
2240 GTTCTAATCTGGTAGACTGTCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299
2261 GTTCTAATCTGGTAGACTGTCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2320
2300 TGGATAAATGAG 2359
2321 TGGATAAATGAG 2380
2360 ATGATCTTGGTAAATTTAAATGTGGCTGATCTCTCTCCCTGCTGAGAAATGAAGCACTG 2419
2381 ATGATCTTGGTAAATTTAAATGTGGCTGATCTCTCTCCCTGCTGAGAAATGAAGCACTG 2440
2420 GCCTGGCTTTGTTCTCCACCTCTTGTCTCCATCAGAGGTCCATTTGTTCCAGTGGATG 2479
2441 GCCTGGCTTTGTTCTCCACCTCTTGTCTCCATCAGAGGTCCATTTGTTCCAGTGGATG 2500
2480 CAAAGAGGCCATCTCTTGAAG 2539
2501 CAAAGAGGCCATCTCTTGAAG 2560
2540 TTGGAGCTTCTCGAGATTTATTTCCACCAAGGAGATTTCCAGGTCCACACCTGCTCCAT 2599
2561 TTGGAGCTTCTCGAGATTTATTTCCACCAAGGAGATTTCCAGGTCCACACCTGCTCCAT 2620

2600 TTGCAATGAGAAATGCTATCCACGAGGGGTTTTCTCTTACCTTCCCTCCCAAGACCTG 2659
2621 TTGCAATGAGAAATGCTATCCACGAGGGGTTTTCTCTTACCTTCCCTCCCAAGACCTG 2680
2660 GATTTTTCCCTCCCAAG 2719
2681 GATTTTTCCCTCCCAAG 2740
2720 CACCTTCAATGAG 2779
2741 CACCTTCAATGAG 2800
2780 TTTGCTCTCTCTTCAAAAGTAATTTTGAAGTCTGATCTCAATTTTCAAGTTAAGTCTGTTA 2839
2801 TTTGCTCTCTCTTCAAAAGTAATTTTGAAGTCTGATCTCAATTTTCAAGTTAAGTCTGTTA 2860
2840 CTTAAGTGATTAACATTTTCTCAATTTGAAGTCTGATCTCAATTTTCAAGTTAAGTCTGTTA 2899
2861 CTTAAGTGATTAACATTTTCTCAATTTGAAGTCTGATCTCAATTTTCAAGTTAAGTCTGTTA 2920
2900 TATTTTGAATTAAGAGATGATTTAAATATGAATCTTATGAGTAAATTTTCAATTTTAT 2959
2921 TATTTTGAATTAAGAGATGATTTAAATATGAATCTTATGAGTAAATTTTCAATTTTAT 2980
2960 TTTAGAGGTATTAACATTTTCAATTTTGAATCTTATGAGTAAATTTTCAATTTTAT 3019
2981 TTTAGAGGTATTAACATTTTCAATTTTGAATCTTATGAGTAAATTTTCAATTTTAT 3040
3020 TTTTATATATGATTAATCTTTTCAAGTGGGAGGCTTTAAATTTTCAAGTCTGTTTATG 3079
3041 TTTTATATATGATTAATCTTTTCAAGTGGGAGGCTTTAAATTTTCAAGTCTGTTTATG 3100
3080 CCAAGAGAGTATTTTACTGTTTGGACAAATGGAAGTAACTTTATGCTTAAATAA 3139
3101 CCAAGAGAGTATTTTACTGTTTGGACAAATGGAAGTAACTTTATGCTTAAATAA 3160
3140 ATTATAGTTGATTTAAAAA 3158
3161 ATTATAGTTGATTTAAGA 3179

RESULT 12
US-09-770-175-6675
; Sequence 6675, Application US/09770175
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2058-001
; CURRENT APPLICATION NUMBER: US/09/770,175
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 8967
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6675
; LENGTH: 3910
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-770-175-6675

Query Match 87.7%; Score 2773.8; DB 32; Length 3910;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2927; Conservative 0; Mismatches 122; Indels 30; Gaps 5;

90 CTAAGGCGGCGGCTCAGACCCAGCGCTCCCTCAGAGTGTAAAGTGAACAGAGGCGC 149
107 CTTGTCGGGTCCGGTTCGGACCTCGCTCGGATGTAAAGTATAACAGAGGGTC 166
150 AGGGAGGTGGTGGGGACAAATGGCCCTGTGAGCCCTGTGGGTCCCGCTTCCCG 209
167 GGGATGGG-----CAGCGTAGGCTGTGAGGCTCGCGGTCGCCCTGTCCCCAG 216

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 05:57:51 ; Search time 276 Seconds

(without alignments)
7705.114 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttcaatccaggaagaag.....agtgatttcaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1375775 seqs, 336063760 residues

Total number of hits satisfying chosen parameters: 2751550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA.New.*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2880.4	91.0	3108	US-10-486-020-51	Sequence 51, Appl
2	2515.2	79.5	2928	US-10-486-020-38	Sequence 38, Appl
3	201	6.4	260	US-10-793-479-29791	Sequence 29791, A
4	135.6	4.3	251	US-10-793-479-2603	Sequence 2603, Ap
5	133.6	4.2	140	US-10-793-479-27286	Sequence 27286, A
6	129.2	4.1	14393	US-10-021-698A-3524	Sequence 3524, Ap
7	64.8	2.0	265	US-10-793-479-34761	Sequence 34761, A
8	60.2	1.9	141646	US-60-550-051-3063	Sequence 3063, Ap
9	58.6	1.9	1000	US-10-021-698A-291	Sequence 291, App
10	55.6	1.8	536	US-10-779-543-17721	Sequence 17721, A
11	55	1.7	27890	US-60-548-091-5662	Sequence 5662, A
12	55	1.7	126872	US-60-548-091-5703	Sequence 5703, Ap
13	54.4	1.7	1035	US-10-779-543-8023	Sequence 8023, Ap
14	50.8	1.6	2637	US-10-724-972A-995	Sequence 995, App
15	50.8	1.6	327636	US-60-550-051-2294	Sequence 2294, Ap
16	50.4	1.6	201	US-60-550-051-21273	Sequence 21273, A
17	50.4	1.6	13485	US-60-550-051-3088	Sequence 3088, Ap
18	49.8	1.6	2100	US-10-416-330-17	Sequence 17, Appl
19	49.6	1.6	1450	US-10-779-543-7878	Sequence 7878, Ap
20	49.2	1.6	5917	US-10-808-727-9	Sequence 9, Appl
21	48.8	1.5	30612	US-10-724-972A-326	Sequence 326, App
22	48.6	1.5	201	US-60-550-051-20553	Sequence 20553, A
23	48.6	1.5	27890	US-60-548-091-5662	Sequence 5662, Ap
24	48.6	1.5	126872	US-60-548-091-5703	Sequence 5703, Ap
25	48.4	1.5	2226	PCT-US04-05654-1714	Sequence 1714, Ap
26	48	1.5	500	PCT-US04-02000-562	Sequence 562, App

27 47.4 1.5 2564 6 US-10-416-330-12
28 47.4 1.5 170546 6 US-10-767-471-10703
C 29 47 1.5 41665 6 US-10-767-471-10744
C 30 46.6 1.5 1035 6 US-10-779-543-4437
C 31 46.6 1.5 1102 6 US-10-767-795-3060
C 32 46.4 1.5 1423 6 US-10-767-795-3350
C 33 46.4 1.5 3053 6 US-10-491-213-108
C 34 46.2 1.5 760 6 US-10-021-698A-161
C 35 46.2 1.5 780 6 US-10-021-698A-540
C 36 46 1.5 885 6 US-10-767-701-20477
C 37 46 1.5 40000 6 US-10-796-280-12600
C 38 46 1.5 107450 6 US-10-796-307-8822
C 39 45.8 1.4 173995 6 US-10-796-307-8746
C 40 45.6 1.4 3636 6 US-10-724-972A-1890
C 41 45.6 1.4 107191 7 US-60-550-051-2971
C 42 45.4 1.4 198285 6 US-10-775-163-338
C 43 45.2 1.4 201 7 US-60-550-051-20567
C 44 45.2 1.4 3885 6 US-10-377-636-3
C 45 45 1.4 481 6 US-10-767-701-17770

ALIGNMENTS

RESULT 1

US-10-486-020-51
; Sequence 51, Application US/10486020
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BARROSO, Ines;
; APPLICANT: BAUGHN, Marian R.; BECHA, Shanya D.;
; APPLICANT: BORSKY, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRAUL, Richard C.;
; APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: HARELIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; REDDY, Roopa;
; APPLICANT: RICHARDON, Thomas W.; SPRAGUE, William W.;
; APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 USN
; CURRENT APPLICATION NUMBER: US/10/486,020
; PRIOR FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: PCT/US02/25465
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 3108


```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7506167CB1
US-10-486-020-51

Query Match      91.0%; Score 2880.4; DB 6; Length 3108;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 3023; Conservative      6; Mismatches      129; Gaps      1;

QY 1 AGGTTTATCCATGAAGAGACAGCAATTTTAAAGTGTAATTCACCAAAAAATAAGCTT 60
DB 72 AGGTTTATCCATGAAGAGACAGCAATTTTAAAGTGTAATTCACCAAAAAATAAGCTT 131
QY 61 CAAATATGTGATGTGAAACTGCCAGAACTAAGGCGGCGCGGCTCAGACCAGCGCTGC 120
DB 132 CAAATATGTGATGTGAAACTGCCAGAACTAAGGCGGCGCGGCTCAGACCAGCGCTGC 191
QY 121 CTCAGAGTGAAGTGTACAGAGAGCGCCAGGGAGGTGGTGGGGACAACATGGGCTTG 180
DB 192 CTCAGAGTGAAGTGTACAGAGAGCGCCAGGGAGGTGGTGGGGACAACATGGGCTTG 251
QY 181 TGAGGCGCTGGGCTGCCCGGTTCCCGAGCTCCCGCCGAGCCCGCTCCACAGTGGTCCG 240
DB 252 TGAGGCGCTGGGCTGCCCGGTTCCCGAGCTCCCGCCGAGCCCGCTCCACAGTGGTCCG 311
QY 241 CTCGGTGGTGTGACGTCGCGCATTCGGGTTCAGACCCAGAGGCTGGTGTCTCCACC 300
DB 312 CTCGGTGGTGTGACGTCGCGCATTCGGGTTCAGACCCAGAGGCTGGTGTCTCCACC 371
QY 301 GCTTGTGTGGCCAGTGTACTCGGTGACCCGAGCAGCAGCTGACGCTATGAGGAG 360
DB 372 GCTTGTGTGGCCAGTGTACTCGGTGACCCGAGCAGCAGCTGACGCTATGAGGAG 431
QY 361 CTTGTGTGTAACCCCTCAGCCCTACCTGGGGTGGTCTCGAGAGAGCTACGACAGTGTG 420
DB 432 CTTGTGTGTAACCCCTCAGCCCTACCTGGGGTGGTCTCGAGAGAGCTACGACAGTGTG 491
QY 421 GCAGCACTACCTGAGAGTATGACACAGATGAGATCCTTATGTTTTCATCGGAAGT 480
DB 492 GCAGCACTACCTGAGAGTATGACACAGATGAGATCCTTATGTTTTCATCGGAAGT 551
QY 481 GTGGTATGTCAGCTGTATGGAATTTTGTGTTTCTCCTTTTGTGGAGAGTTT 540
DB 552 GTGGTATGTCAGCTGTATGGAATTTTGTGTTTCTCCTTTTGTGGAGAGTTT 611
QY 541 AGATCGTTAGGAGTGGCTTTACGTGGGAGAGACAAACTTGGTCCACGCTTCT 600
DB 612 AGATCGTTAGGAGTGGCTTTACGTGGGAGAGACAAACTTGGTCCACGCTTCT 671
QY 601 GGAATAATTGAAGAAAATGTAACTACTTGAATAATTTAGCTTATTCAAAAAGAGTAT 660
DB 672 GGAATAATTGAAGAAAATGTAACTACTTGAATAATTTAGCTTATTCAAAAAGAGTAT 731
QY 661 GAGGCTATGAGTATGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 720
DB 732 GAGGCTATGAGTATGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 791
QY 721 GAAGCAGAGAGTTGGAGGCAACTGTGAAAAGCTGAAAGGTCCCAATTTCTGAATTTGAG 780
DB 792 GAAGCAGAGAGTTGGAGGCAACTGTGAAAAGCTGAAAGGTCCCAATTTCTGAATTTGAG 851
QY 781 GATGAATCTCTGTCTAGAAAAGACTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 840
DB 852 GATGAATCTCTGTCTAGAAAAGACTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 911
QY 841 GATGAATTTGATGGCGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
DB 912 GATGAATTTGATGGCGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 971
QY 901 CTCAATATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTTAAATGATGTAAGAA 960

```

```

DB 972 CTCAAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 1031
QY 961 CGACGGCTATAGCAATATAAAGATGCTTTGAAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1020
DB 1032 CGACGGCTATAGCAATATAAAGATGCTTTGAAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1091
QY 1021 AAAACAGCTTTTTCAGCAAGAGAGCTGAAGTATGAAAGGAGAGAGTGAAGTGAACCTTAATAAA 1080
DB 1092 AAACAGCTTTTTCAGCAAGAGAGCTGAAGTATGAAAGGAGAGAGTGAAGTGAACCTTAATAAA 1151
QY 1081 CAGAAAATACATTTGAAGACTCCAAAGTACACGCGAGAACAGTTCGAAATGATTAAGAA 1140
DB 1152 CAGAAAATACATTTGAAGACTCCAAAGTACACGCGAGAACAGTTCGAAATGATTAAGAA 1211
QY 1141 AATCACATCAAGACCTTGACTGGACACTTCCCAATGATGAAAGATCAGGCTGCTGCTT 1200
DB 1212 AATCACATCAAGACCTTGACTGGACACTTCCCAATGATGAAAGATCAGGCTGCTGCTT 1271
QY 1201 GAAGAAACACACACCGATGATGATTAACCTGGATTTAGAAAGTGAACAGTGAATCGGAAAT 1260
DB 1272 GAAGAAACACACACCGATGATGATTAACCTGGATTTAGAAAGTGAACAGTGAATCGGAAAT 1331
QY 1261 GGTGCTTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTTCATGCTGCTAAG 1320
DB 1332 GGTGCTTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTTCATGCTGCTAAG 1391
QY 1321 TTAATGCTTTTAAACCTTAGAAGGAGAGAGAAACCAATTTATATTCAGTTGCTT 1380
DB 1392 TTAATGCTTTTAAACCTTAGAAGGAGAGAGAAACCAATTTATATTCAGTTGCTT 1451
QY 1381 GAAGCTTCATATAACCAAGAGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1440
DB 1452 GAAGCTTCATATAACCAAGAGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1511
QY 1441 GCATCTTTGCGAGTCAGAAAAACACATTTTGAATAAGAGAAATCAGAGCTTCAACAGAAA 1500
DB 1512 GCATCTTTGCGAGTCAGAAAAACACATTTTGAATAAGAGAAATCAGAGCTTCAACAGAAA 1571
QY 1501 CTTAAAGTAAATGACTGAATTTATCAAGAAAATGAATAAACTCCACAGAGAAATTAACA 1560
DB 1572 CTTAAAGTAAATGACTGAATTTATCAAGAAAATGAATAAACTCCACAGAGAAATTAACA 1631
QY 1561 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAAACTTTCTAAAGTATGATGAAAAGATC 1620
DB 1632 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAAACTTTCTAAAGTATGATGAAAAGATC 1691
QY 1621 AGCCATGCCACTGAGAGCTGGAGCCTATAGAAAGCGAGCCCAAGATCTTGAAGAGAA 1680
DB 1692 AGCCATGCCACTGAGAGCTGGAGCCTATAGAAAGCGAGCCCAAGATCTTGAAGAGAA 1751
QY 1681 TTGAGAGAACTATTTCATTTCTTATCAAGGGCAGATTTTCCCATGAGAAAAAGCAGAT 1740
DB 1752 TTGAGAGAACTATTTCATTTCTTATCAAGGGCAGATTTTCCCATGAGAAAAAGCAGAT 1811
QY 1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAACTCTCAATGTTTAAAGGAGAGAAAT 1800
DB 1812 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAACTCTCAATGTTTAAAGGAGAGAAAT 1871
QY 1801 GCTCACACACAGACAAAAATTAAGTGAACACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
DB 1872 GCTCACACACAGACAAAAATTAAGTGAACACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1931
QY 1861 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTTGCC 1920
DB 1932 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTTGCC 1967
QY 1921 TCACCAATTTGGGTTGGCCTTTCATCTGAAAAAAGAGCTTTTCTCTCTCTCTCCCACTTTG 1980
DB 1968 ----- 1967
QY 1981 GAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGGCTCAGAGGC 2040
DB 1968 -----AGAGGCTCAGAGGC 1982

```


	3121	TAACTTTATGCTTAAATAAATTATAGTTGATTAAAAA	3150
Qy			
	3063	TAACTTTATGCTTAAATAAATTATAGTTGATTAAAAA	3100
Db			
RESULT 2			
US-10-486-020-38			
; Sequence 38, Application US/10486020			
; GENERAL INFORMATION:			
; APPLICANT: AZIMZAL, Yalda; BARROSO, Ines;			
; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;			
; APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;			
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;			
; APPLICANT: FORTSYTHE, Ian J.; GIERZEN, Kimberly J.;			
; APPLICANT: GOSEVAD, Ann E.; GRAUL, Richard C.;			
; APPLICANT: GRIFPIN, Jennifer A.; GURURAJAN, Rajagopal			
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;			
; APPLICANT: KABLES, Amy E.; KHAN, Farrah A.;			
; APPLICANT: LEE, Sally;			
; APPLICANT: LI, Joana X.; REDDY, Roopa;			
; APPLICANT: RICHARDON, Thomas W.; SPRAGUE, William W.;			
; APPLICANT: SWARNAKAR, Anica; TANG, Y. Tom;			
; APPLICANT: WARREN, Bridget A.; XU, Yuming;			
; APPLICANT: YAO, Monique G.; YUE, Henry;			
; APPLICANT: YUE, Hubbin			
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH			
; FILE REFERENCE: PF-1126 USN			
; CURRENT APPLICATION NUMBER: US/10/486,020			
; PRIOR FILING DATE: 2004-02-05			
; CURRENT APPLICATION NUMBER: PCT/US02/25465			
; PRIOR FILING DATE: 2002-08-02			
; PRIOR APPLICATION NUMBER: US 60/311,017			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/313,070			
; PRIOR FILING DATE: 2001-08-17			
; PRIOR APPLICATION NUMBER: US 60/313,071			
; PRIOR FILING DATE: 2001-08-17			
; PRIOR APPLICATION NUMBER: US 60/314,678			
; PRIOR FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: US 60/316,592			
; PRIOR FILING DATE: 2001-08-31			
; PRIOR APPLICATION NUMBER: US 60/317,913			
; PRIOR FILING DATE: 2001-09-07			
; PRIOR APPLICATION NUMBER: US 60/322,182			
; PRIOR FILING DATE: 2001-09-14			
; PRIOR APPLICATION NUMBER: US 60/340,747			
; PRIOR FILING DATE: 2001-12-07			
; PRIOR APPLICATION NUMBER: US 60/342,761			
; PRIOR FILING DATE: 2001-12-20			
; Remaining Prior Application data removed - See File Wh			
; NUMBER OF SEQ ID NOS: 54			
; SOFTWARE: PERL Program			
; SEQ ID NO 38			
; LENGTH: 2928			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; NAME/KEY: misc feature			
; FEATURE:			
; OTHER INFORMATION: Incyte ID No: 6715627CBI			
US-10-486-020-38			
Qy	90	CTAAGGGGGCGGGGCTCAGACCAAGCGCTGCCTCAGGATGTGA	
Db	35	CTTGTCGGGCTCGGGCTCGGACTCGCTGCCTCGGATGTGA	
Qy	150	AGGGAGAGTGTGGGGAGACAACATGGCCCTGTGAGGCCCTGTG	
Db	95	GCGATGTGTACGCTGAGCCCTGTGTAGCCCTGT-----CGGG	
Query Match 79.5%; Score 2515.2; DB 6; 1			
Best Local Similarity 93.6%; Pred.No. 0; 158; Matches			
Matches 2708; Conservative 0; Mismatches			

